

DEPARTEMEN STATISTIKA, INSTITUT PERTANIAN BOGOR

PEMODELAN
Additive Main-effect
& Multiplicative
Interaction (AMMI):
KINI DAN YANG
AKAN DATANG

Buku Riset

2011

KAMPUS IPB DARMAGA, JL. MERANTI W22 LV 3

PEMODELAN AMMI: KINI DAN YANG AKAN DATANG
BUKU RISET 2011

Editor :

Ahmad Ansori Mattjik

I Made Sumertajaya

Alfian Futuhul Hadi

Gusti Ngurah Adhi Wibawa

Kontributor:

A A Mattjik, I Made Sumertajaya, Aunuddin,
Totong Martono, Khairil Anwar Notodiputro, Agus M. Sholeh
Rusida Yulianti, I Nyoman Mindra Jaya, Pika Sivianti,
Pepi Novianti, I Gusti Ngurah Adhi Wibawa,
Halimatus Sa'diyah, Sri Zulhayana,
Alfian Futuhul Hadi

Hak Cipta dilindungi Undang-Undang

Dilarang mengutip atau memperbanyak sebagian atau
seluruh isi buku teks ini tanpa ijin tertulis dari Penerbit

Diterbitkan oleh IPB PRESS

Edisi ke-pertama: September 2011

ISBN 978-602-96772-6-3

PENGANTAR

Alhamdulillah, puji sukur pada Allah SWT, dengan rahmatNya telah tersusun buku hasil penelitian tentang interaksi genetik dan lingkungan (IGL) dengan menggunakan pendekatan model aditif untuk pengaruh utama genetik dan lingkungan serta interaksinya multifikatif.

Perhatian terhadap IGL dipicu oleh banyaknya penelitian yang melibatkan banyak lokasi pada berbagai musim tanam, apabila dianalisis dengan model aditif saja menghasilkan keragaman percobaan yang cukup besar. Kenyataan ini sering menyulitkan untuk memberikan interpretasi dan memberikan rekomendasi atas genotipa yang dicobakan.

Dimulai dengan menyertakan struktur peragam spasial (*spatial covariance*) kedalam model, kemudian dikembangkan menjadi model genetik sebagai pengaruh random analog dari model efek utama aditif dan interaksi multiplikasi atau yang kemudian dikenal dengan model *Additive Main-effects and Multiplicative Interaction* (AMMI). Keunikan model multiplikatif AMMI terletak pada visualisasi interaksi melalui biplot hasil dari penguraian singular dekomposisi. Model AMMI dibangun berlandaskan pada distribusi data Normal (gaussian), juga teknik komputasi yang sederhana sehingga dengan cepat dapat digunakan secara luas. Dengan alasan itu pilihan untuk mengembangkan model AMMI lebih rinci menjadi dapat dipertanggungjawabkan. Diharapkan hasil penelitian ini dapat memberikan sedikit kontribusi pada kemajuan ilmu pengetahuan, khususnya pada ilmu Statistika dan Pemuliaan Tanaman atau pada bidang lain yang lebih luas nantinya.

Utamanya penelitian ini melibatkan empat staf Departemen Statistika, satu orang staf Balai Penelitian, dua mahasiswa program Doktor, tujuh mahasiswa program Master, dengan ikutan tiga orang mahasiswa Sarjana. Sampai akhir pelaporan, bulan November 2010, hampir semua mahasiswa telah selesai, satu orang mahasiswa S2 selesai Desemberr 2010, kecuali kedua mahasiswa program Doktor yang baru akan selesai pada bulan Desember 2011. Sudah dihasilkan 12 artikel ilmiah, 5 diantaranya pada forum seminar internasional, dan 7 artikel diterbitkan pada jurnal nasional, disamping laporan penelitian dan disertasi.

Tidak terelakkan lagi penelitian ini berjalan karena dukungan dan kerjasama berbagai pihak. Karenanya kami sampaikan ucapan terimakasih yang sedalam-dalamnya kepada:

- Kementerian Pendidikan Nasional Republik Indonesia melalui Dirjen Pendidikan Tinggi yang telah mendukung penelitian ini dengan menyediakan dana seperti pada kontrak Hibah Pasca Sarjana Nomor 266/13.11/PL/2008, 41/13.24.4/SPK/BG-PD/2009, dan 1/13.24.4/PD/2010.

Model AMMI Kini dan Yang Akan Datang

- Rektor melalui Lembaga Penelitian dan Pengabdian Kepada Masyarakat dan Sekolah Pasca Sarjana IPB yang mendukung penyelesaian dan mengevaluasi penelitian hibah ini.
- Kementerian Pertanian Republik Indonesia melalui Balai Penelitian Tanaman Padi Sukamandi dan Balai Besar Penelitian Tanaman Pangan Cimanggu yang telah membantu menyediakan data percobaan multi lokasi.
- Serta semua pihak, dengan permohonan maaf tidak dapat disebutkan satu persatu, yang telah membantu menyelesaikan penelitian ini.

Dari hasil penelitian diharapkan muncul ide baru dalam mengembangkan model seleksi daya adaptasi tanaman yang efektif dan efisien yang perlu di lakukan penelitian lebih lanjut untuk dapat membantu para pemulia tanaman dalam mencari genotipe unggulan.

Bogor, Agustus 2011

A. A. Matjjik

DAFTAR ISI

PENGANTAR	3
DAFTAR ISI	5
BAB I. PENDAHULUAN	8
1.1 Motivasi	8
1.2 Analisis Statistik Interaksi Genotipe Lingkungan: Peningkatan Akurasi	9
1.3 AMMI dan Analisis IGL.....	10
BAB II. PENGEMBANGAN MODEL AMMI: MASALAH DAN PENYELESAIANNYA	14
2.1 Pengembangan Model AMMI.....	14
2.2 Roadmap Riset 2008-2010.....	16
BAB III. ANALISIS STABILITAS AMMI: DARI INDEKS STABILITAS AMMI SAMPAI RESAMPLING BOOTSTAP	20
3.1 Indeks Stabilitas AMMI	20
3.2 Uji Lanjut Pengaruh Interaksi pada Model AMMI	21
3.3 Pendugaan Kestabilan Genotipe pada Model AMMI menggunakan Resampling Bootstrap.....	22
3.4 Indeks Stabilitas AMMI untuk Penentuan Stabilitas Genotipe pada Percobaan Multilokasi (Halimatus Sa'diyah dan Ahmad Ansori Mattjik)	26
3.5 PENGUJIAN SUBHIPOTESIS PENGARUH INTERAKSI GENOTIPE × LINGKUNGAN PADA MODEL AMMI (A A Mattjik, Rusida Yuliyanti dan Totong Martono)	38
3.6 Pendugaan Daerah Kepercayaan dan Kestabilan Genotipe Pada Model AMMI Menggunakan Metode Resampling Bootstrap (Pepi Novianti, A A Mattjik & I M Sumertajaya)	48
BAB IV. Penanganan Data Hilang pada Model AMMI	62
4.1 EM-AMMI	63
4.2 Metode g-Connected.....	63

4.3	PENDUGAAN DATA TIDAK LENGKAP DALAM AMMI (I Made Sumertajaya, G N A Wibawa, Journal BIAStatistics (2008) Vol. 2, No. 2, hal. 58-71)	66
BAB V.	Model AMMI untuk Respon Ganda	82
5.1	Evaluasi Tingkat Validitas Metode Penggabungan Respon: Indeks Penampilan Tanaman (I Made Sumertajaya, G N A Wibawa, Seminar Nasional Matematika, UNY Yogyakarta, 22 Nopember 2008)	84
BAB VI.	AMMI untuk Model Campuran	101
6.1	PENANGANAN MODEL CAMPURAN PADA PERCOBAAN MULTILOKASI MENGGUNAKAN BLUP (Halimatus Sa'diyah, A A Mattjik, dan I M Sumertajaya)	103
BAB VII.	Penggunaan Model Persamaan Sruktural pada Model AMMI Igl	116
7.1	ANALISIS INTERAKSI GENOTIPE X LINGKUNGAN MENGGUNAKAN STRUCTURAL EQUATION MODELING (I M Sumertajaya, A A Mattjik, I G N M Jaya, Journal Pythagoras Vol. 4, No. 1, hal. 15-32)	120
7.2	ANALISIS INTERAKSI GENOTIPE-LINGKUNGAN MENGGUNAKAN PARTIAL LEAST SQUARE PATH MODELING (I Made Sumertajaya, Jurnal, BIAStatistika Vol. 4, No. 1, hal. 9-18)	138
BAB VIII.	Pengamatan Pencilan pada Analisis Interksi Genotipe Lingkungan ...	154
8.1	Stabilitas Non Parametrik	155
8.2	Robust-AMMI dengan Model Faktor Analitik	156
8.3	Analisis Stabilitas Genotipe Padi dengan Indeks Stabilitas Nonparametrik Thennarasu (S. Zulhayana, I M Sumertajaya, A A Mattjik)	160
8.4	Handling Outlier Observation in Two-Ways Table by Robust Alternating Regression of Fanova Models: Towards Robust AMMI Models (Alfian F. Hadi & A. A. Mattjik, Jurnal ILMU DASAR V. 12 No. 2, 2011, in press)	168
BAB IX.	Ketaknormalan Data pada Model AMMI	182
9.1	Pendekatan Transformasi Kenormalan	183
9.2	Model AMMI Terampat	186

9.3	GENERALIZED AMMI MODELS FOR ASSESSING THE ENDURANCE OF SOYBEAN TO LEAF PEST (Alfian F. Hadi , A. A. Mattjik & I M Sumertajaya, Jurnal Ilmu Dasar Terakreditasi. Vol 11. NO. 2. Juli 2010. P 151-159).....	188
9.4	Ketaknormalan Data dan Pengamatan Pencilan Pada Model AMMI (A F Hadi dan A A Mattjik, Simposium dan Kongres PERIPI November 2009).....	202
BAB X.	Komputasi Bayessian pada Model AMMI	219
10.1	Bayessian	219
10.2	Pendekatan Bayesian pada Pendugaan Matriks Interaksi.....	222
10.3	Model AMMI dengan Komputasi Bayessian (G N A Wibawa, Aunuddin, A A Mattjik)	227
10.4	Pendekatan Metode Bayes Untuk Pendugaan Pengaruh Interaksi Pada Model AMMI (Pika Silvianti , Khairil A. Notodiputro, dan I Made Sumertajaya)	235
BAB XI.	PENUTUP.....	244
11.1	Ringkasan.....	245
11.2	Pengembangan Lebih Lanjut (<i>futher works</i>).....	247
DAFTAR PUSTAKA.....		248

PENDAHULUAN

1

PENDAHULUAN

1.1 Motivasi

Sumberdaya manusia merupakan salah satu pilar pembangunan nasional. Keberhasilan pembangunan nasional akan ditentukan oleh kualitas sumberdaya manusia yang dimilikinya. Dengan demikian untuk mencapai tujuan pembangunan nasional, pengembangan kualitas sumber daya manusia sangat perlu dilakukan. Beberapa faktor yang dapat meningkatkan kualitas sumberdaya manusia antara lain pangan, nutrisi dan kesehatan. Oleh karena itu, penyediaan pangan yang memadai merupakan salah satu kunci dalam peningkatan kualitas sumber daya manusia.

Dalam rangka ikut serta membantu meningkatkan kualitas hidup para petani dapat ditempuh melalui berbagai program antara lain menyediakan berbagai varietas unggulan yang memiliki tingkat produksi tinggi, berumur pendek, tahan terhadap serangan hama dan penyakit. Untuk mendapatkan varietas-varietas unggulan perlu dilakukan berbagai penelitian seperti rekayasa genetik, persilangan antar galur, serta uji daya adaptasi berbagai genotipe pada berbagai kondisi lingkungan. Secara umum, penelitian ini ingin berkontribusi dalam pengembangan metode yang berkenaan dengan analisis interaksi yang banyak digunakan dalam kajian daya adaptasi genotipe suatu tanaman pada berbagai kondisi yang tersedia. Menyediakan pemodelan untuk seleksi melalui uji adaptabilitas dan stabilitas yang amat diperlukan untuk mendukung upaya memperoleh varietas unggul.

Untuk menjamin tersedianya pangan yang berkualitas perlu dilakukan berbagai kajian diantaranya meneliti tentang daya adaptasi berbagai genotipe suatu tanaman pada berbagai kondisi yang tersedia, yang sering dikenal dengan sebutan uji lokasi ganda (*multi locations/environments trial*). Uji lokasi ganda melibatkan dua faktor utama yaitu genotipe tanaman dan kondisi lingkungan (lingkungan: tempat (*site*), musim, perlakuan agronomis (*agronomy treatment*)). Dari uji lokasi ganda diharapkan mampu memilah pengaruh utama (genotipe dan lingkungan) dan pengaruh interaksi antara genotipe dengan lingkungan (IGL). Dari pengaruh interaksi tersebut dapat dipilah genotipe-genotipe yang mampu beradaptasi pada berbagai kondisi lingkungan (genotipe stabil) dan genotipe-genotipe yang hanya sesuai pada lingkungan tertentu (genotipe spesifik). Untuk mampu memilah kedua pengaruh ini dengan baik dibutuhkan pendekatan analisis yang tepat.

1.2 Analisis Statistik Interaksi Genotipe Lingkungan: Peningkatan Akurasi

Rekomendasi terhadap genotipe sebagai jenis/varietas tanaman baru untuk tujuan komersial memerlukan prediksi yang handal dan akurat pada rataan produksi dari setiap varietas di berbagai lingkungan serta pengetahuan yang memadai tentang interaksi genotipe dan lingkungan. Informasi ini dapat diperoleh dari beberapa percobaan atau sering disebut *Multi Environment Trial* (MET). Setiap tahun banyak biaya yang dikeluarkan untuk mendapatkan

informasi ini. Untuk memaksimalkan efisiensi biaya dalam pekerjaan ini, sangat penting melakukan analisis data dengan menggunakan metode statistika yang tepat. Tiga aspek kunci dari analisis tersebut adalah asumsi tentang pengaruh varietas dan interaksinya dengan lingkungan serta tentang galat percobaan dari setiap individual percobaan. Beberapa disain percobaan yang banyak dilakukan antara lain rancangan acak kelompok atau rancangan acak lengkap.

Data dari lapang sering memperlihatkan variasi spasial, disebut demikian karena merupakan fungsi dari lokasi dan plot di lapang. Gilmour et al. (1997) menyampaikan suatu metode analisis dimana variasi spasial dimodelkan untuk menghasilkan pendugaan pengaruh perlakuan yang mempunyai akurasi dan presisi yang lebih baik dibandingkan metode tradisional, seperti rancangan acak kelompok dan IB (Gleeson dan Cullis, 1987). Cullis et al (1998) mengadopsi pendekatan tersebut dalam analisis model campuran pada data MET. Hal ini memungkinkan untuk memisahkan struktur peragam spasial (*spatial covariance*) pada setiap percobaan. Langkah ini merupakan tahap utama untuk menghasilkan dugaan yang akurat pada rata-rata varietas (Smith et al., 2001). Smith et al (2001) mengembangkan pendekatan tersebut dengan menggunakan model multiplikatif untuk pengaruh interaksi varietas dan lingkungan. Model ini memungkinkan untuk memisahkan ragam genetik untuk setiap percobaan dan memberikan suatu model parsimonious dan model yang lebih mudah diartikan tentang peragam genetik antar pasangan dalam percobaan. Model genetik dapat sebagai pengaruh random analog dari model efek utama aditif dan interaksi multiplikasi (*Aditif Main effect and Multiplicative Interaction (AMMI)*) (Gauch, 1992).

1.3 AMMI dan Analisis IGL

Untuk mengkaji pengaruh genotipe pada berbagai kondisi lingkungan dapat dilakukan melalui uji multilokasi. Uji ini dapat dilakukan dengan melibatkan berbagai genotipe pada berbagai kondisi lingkungan, yang meliputi tempat, tahun tanam dan berbagai perlakuan agronomi lainnya. Uji ini dilakukan seperti halnya rancangan percobaan biasa, dengan blok disarangkan ke dalam lingkungan. Analisisnya sering disebut analisis ragam gabungan (*Composite Analysis of Variance*). Modelnya dapat dituliskan sebagai berikut:

$$Y_{ijk} = \mu + \beta_{k(j)} + G_i + L_j + GL_{ij} + \varepsilon_{ijk}$$

Permasalahan selanjutnya yang sering dihadapi adalah bagaimana menguraikan pengaruh interaksi genotipe dengan lingkungan secara efektif. Berbagai metode telah dikembangkan oleh berbagai tokoh statistika seperti Eberhart Russel, Finlay Wilkinson, dan Tukey. Metode yang dikemukakan ketiga tokoh tersebut cukup efektif dalam memilah genotip-genotipe yang stabil dan spesifik. Namun pendekatan ini masih meninggalkan keragaman interaksi yang cukup besar yang terjadi karena pendekatan ini hanya menjelaskan komponen linier dari pengaruh interaksi sehingga apabila pola interaksi genotipe terhadap lingkungan tidak linier akan menyisakan keragaman yang cukup besar.

Kelamahan ini menjadi pemicu berkembangnya metode AMMI (*Additive Main Effects and Multiplication Interaction*). Tokoh-tokoh yang mengembangkan metode ini antara lain Gauch and Zobel (1988); Zobel et al, (1988); Crossa et al, (1990). AMMI sangat efektif menjelaskan interaksi genotipe dengan lingkungan. Penguraian pengaruh interaksi dilakukan dengan model bilinear, sehingga kesesuaian tempat tumbuh bagi genotipe akan dapat dipetakan dengan jelas. Terlihat dengan jelas dalam pemetaan genotipe dan lingkungan secara simultan dengan menggunakan biplot.

Ada tiga manfaat utama penggunaan analisis AMMI yaitu pertama analisis AMMI dapat digunakan sebagai analisis pendahuluan untuk mencari model yang lebih tepat. Jika tidak ada satupun komponen yang nyata maka pemodelan cukup dengan pengaruh aditif saja. Sebaliknya jika hanya pengaruh ganda saja yang nyata maka pemodelan sepenuhnya ganda, berarti analisis yang tepat adalah analisis komponen utama saja. Sedangkan jika semua komponen interaksi nyata berarti pengaruh interaksi benar-benar sangat kompleks, tidak memungkinkan dilakukannya pereduksian tanpa kehilangan informasi penting (Bradu and Bariel, 1978; Gauch, 1985).

Manfaat kedua dari analisis AMMI adalah untuk menjelaskan interaksi genotipe dan lingkungan. AMMI dengan biplotnya meringkas pola hubungan antar genotip, antar lingkungan, dan antara genotipe dan lingkungan (Kempton, 1984; Zobel et al, 1988; Crossa, 1990). Kegunaan ketiga adalah meningkatkan keakuratan dugaan respon interaksi genotipe x lingkungan.

Dari beberapa penelitian yang telah dilakukan diketahui bahwa pendekatan AMMI lebih baik dalam mengkaji struktur interaksi antara genotipe dengan lingkungan. Model AMMI mampu menjelaskan interaksi dengan baik melalui

model interaksi lengkap atau dikenal sebagai suku multiplikatif/bilinear (Sumertajaya, 1998). Groenen, P.J.F & A. J. Koning, (2004b) menunjukkan penggunaan biplot pada model bilinear sebagai cara baru memvisualisasi interaksi pada model aditif (ANOVA model). Struktur interaksi diuraikan dari matriks sisaan komponen aditif dengan memanfaatkan sifat matematis penguraian nilai singular (*singular value decomposition*, SVD). SVD merupakan pendekatan kuadrat terkecil dengan reduksi dimensi (pangkat matriks) data yang terbaik dan menyediakan penyajian secara grafis yang dikenal secara luas dengan nama Biplot.

Kekuatan eksplorasi model multiplikatif AMMI terletak pada visualisasi interaksi melalui biplot ini. Selain itu model AMMI dibangun dengan landasan teori pemodelan pada distribusi data Normal (gaussian) yang mapan, teknik komputasi yang sederhana dan karenanya ia telah secara luas digunakan. Alasan-alasan tersebut membuat pilihan untuk mengembangkan model AMMI lebih lanjut menjadi rasional, sehingga kami berupaya melakukannya guna memberikan sedikit saja kontribusi pada kemajuan ilmu pengetahuan khususnya dalam Statistika dan Pemuliaan Tanaman atau pada bidang lain yang lebih luas nantinya.

Halaman ini sengaja dikosongkan

2 **PENGEMBANGAN MODEL AMMI: MASALAH DAN PENYELESAIANNYA**

2.1 Pengembangan Model AMMI

Kebutuhan pemodelan untuk seleksi melalui uji adaptabilitas dan stabilitas amat diperlukan untuk mendukung upaya memperoleh varietas yang unggul. Kebutuhan ini semakin nyata untuk menghindari terjadinya kesalahan pelepasan varietas yang mungkin dapat merugikan petani dan menimbulkan kekacauan/kepanikan karena kelangkaan pangan akibat prediksi pemenuhan pangan nasional yang salah.

Seiring dengan permasalahan riil pada pemuliaan tanaman pangan, beberapa hal dari pendekatan AMMI perlu dikembangkan untuk memperluas cakupan analisis. Hal ini dimaksudkan sebagai upaya optimalisasi penggunaan model AMMI dalam mendekomposisi IGL rangka pemenuhan pangan umat manusia. Melalui penelitian ini akan diperoleh pendekatan pemodelan multiplikatif yang

komprehensif dan terintegrasi dalam upaya menjamin tersedianya pangan melalui diperolehnya varietas tanaman pangan yang unggul dan mempunyai daya adaptasi pada berbagai kondisi lingkungan Indonesia.

Beberapa upaya pengembangan dapat disebutkan disini antara lain, pertama tentang metode komputasi bayesian. Penanganan keacakan model AMMI dapat dilakukan dengan pendekatan Bayes pada pendugaan matriks interaksi. Meskipun baru dilakukan pada kondisi keragaman data yang sangat terbatas, namun telah diperoleh informasi awal bahwa pendekatan Bayes selalu memiliki MSE yang lebih kecil dari metode kuadrat terkecil (MKT). Kedua, pada model campuran pada AMMI dapat dilakukan dengan pendekatan BLUP untuk pendugaan matriks interaksi. Meski belum diperoleh perbandingan dengan metode kuadrat terkecil namun diyakini penggunaan BLUP berada pada asumsi (keheterogenan ragam) yang tepat. Kedua hasil ini menyisakan pertanyaan tentang (i) perbandingan performa keduanya, (BLUP dan BAYES) pada pendugaan matriks interaksi, dan (ii) karena penggunaan Bayes yang telah dilakukan baru pada kondisi keragaman data yang sangat terbatas, maka perlu dilakukan pengembangan metode Bayes pada kondisi keragaman data yang lebih umum. Disamping itu perlu pula pengembangan Bayesian pada pemodelan AMMI secara lengkap dalam mendekomposisi interaksi bukan hanya pada pendugaan matriks interaksi.

Sementara itu pada mengembangkan metode robust, telah diperoleh model multiplikatif melalui algoritma regresi bolak-balik. Disamping itu dalam model ini pula telah diperoleh sekaligus identifikasi adanya pencilan didalam data tabel dua arah atau pun pada matriks interaksi yang akan didekomposisi. Sedangkan pada sisi pengembangan metode non parametrik, telah pula diperoleh analisis kestabilan non-parametrik yang secara teoritik diperkirakan akan pula memiliki sifat kekekaran terhadap pencilan, meskipun belum diperoleh hasilnya pada saat ini. Kedua hasil ini menggugah kami untuk menuntaskan pertanyaan tentang bagaimana performa keduanya terhadap suatu kondisi terdapatnya pengamatan pencilan. Disisi lain, masalah yang telah sejak awal penelitian dipikirkan, yaitu tentang pengembangan model AMMI pada pengamatan kualitatif/kategorik yang berdistribusi non-normal telah pula diperoleh informasi awal bahwa pendugaannya dilakukan dengan memanfaatkan algoritma regresi bolak-balik. Yang ini berarti terdapat keparalelan dengan model robust.

2.2 Roadmap Riset 2008-2010

Pada dasawarsa terakhir metode seleksi adaptabilitas genotype melalui percobaan multilokasi telah banyak menggunakan model AMMI. Sampai saat ini model AMMI telah berkembang baik untuk data kuantitatif (sebaran normal), sebut saja model AMMI, maupun untuk data kategorik (sebaran bukan normal) yang disebut G-AMMI.

Model AMMI pada dasarnya adalah model dengan faktor tetap tetap (fixed model) yang mengasumsikan genotype dan lingkungan ditentukan secara subyektif oleh peneliti dan kesimpulan yang diharapkan hanya terbatas pada genotype dan lingkungan yang dicobakan saja. Namun telah pula berkembang model campuran, campuran antara faktor tetap dan acak, (Mixed/Random AMMI). Model ini memperluas cakupan kesimpulan, dimana lingkungan bersifat acak dan kesimpulan untuk faktor lingkungan berlaku untuk populasi lingkungan dalam hal ini lokasi budidaya tanaman di seluruh Indonesia. Semetara itu, untuk data kualitatif telah berkembang pula model kategorik (GLM-AMMI/Generalized Linear Model AMMI) yaitu jika respon yang diamati bersifat kategorik seperti banyaknya serangan hama (sebaran poisson) dan data sebaran binomial dalam bentuk persentase (Eeuwijk, 1995; Hadi, Sumertajaya & Mattjik, 2005; serta Hadi, 2006).

Suatu percobaan multilokasi membutuhkan biaya yang tidak sedikit, disamping melibatkan begitu banyak genotype, ia juga dilakukan pada berbagai lingkungan budidaya di Indonesia. Desain percobaan di tiap-tiap lokasi haruslah sama satu sama lain. Beberapa asumsi dan kaidah perancangan percobaan adakalanya tidak dapat dipenuhi, seperti kebebasan antar pengamatan, termasuk didalamnya asumsi sebaran data. Beberapa masalah akan muncul dalam analisis. Pada penelitian tahun kedua ini masalah model seleksi adaptabilitas melalui model AMMI dikhususkan pada pengembangan metode pendugaan model AMMI pada berbagai kondisi.

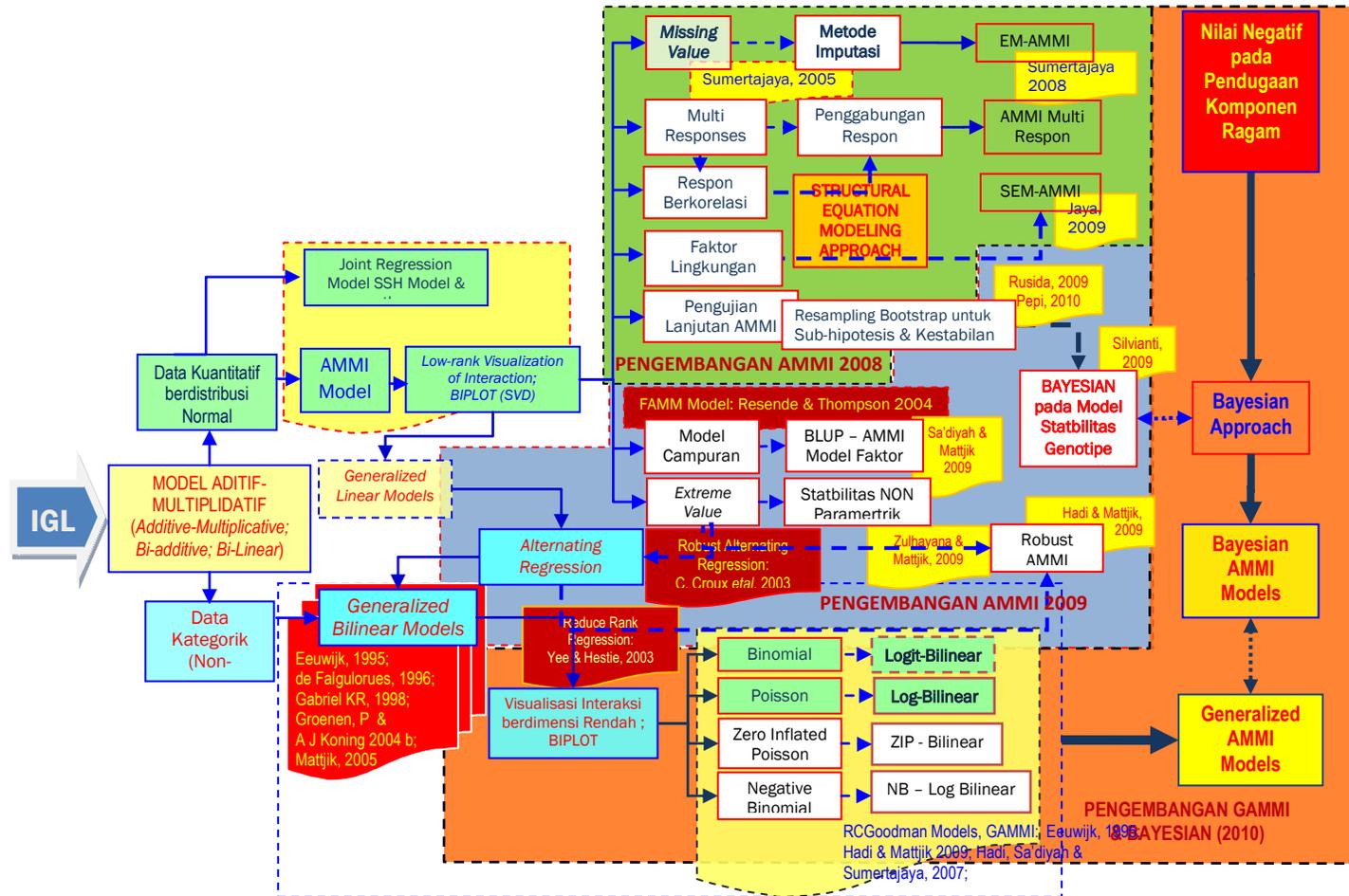
Upaya pengembangan model AMMI ini dituangkan dalam sebuah roadmap riset seperti pada Gambar 1. Beberapa hal terkait pengembangan model AMMI selama kurun waktu 2008-2011 antara lain adalah :

1. Sejauh ini model AMMI telah mampu menerangkan pengaruh interaksi genotype \times lingkungan. Pengembangan juga dilakukan dengan memformulasi index stabilitas dari biplot AMMI. Namun demikian, model AMMI belum dapat digunakan sebagai penguji lanjut (post-hoc) untuk menguji kontribusi yang diberikan oleh genotype dan lingkungan terhadap pengaruh interaksi tersebut. Nilai kriteria uji dari pengujian ini belum diketahui. Pengujian sub-

- hipotesis pada interaksi genotipe \times lingkungan dilakukan dengan aproksimasi menggunakan resampling data dengan pengembalian.
2. Interpretasi hasil AMMI dilihat dari Biplot yang memvisualisasikan skor komponen utama interaksi pertama dan komponen utama interaksi kedua ke dalam grafik berdimensi dua. Biplot AMMI hanyalah suatu analisis eksplorasi dan tidak menyediakan pengujian hipotesis, sehingga interpretasi yang dihasilkan dapat berbeda bagi setiap peneliti. Suatu metode secara inferensia untuk interpretasi hasil biplot AMMI dikembangkan melalui penggunaan resampling bootstrap.
 3. Permasalahan yang juga sering terjadi dilapangan yaitu genotipe yang dicobakan tidak sama pada semua lingkungan. Ada sebagian genotipe yang sama pada setiap lingkungan dan sebagian lagi genotipe yang dicobakan tidak sama. Struktur data seperti ini tentunya membutuhkan penanganan tersendiri agar dapat dianalisis dengan menggunakan model AMMI. Sumertajaya, 2006 melaporkan penggunaan *g-connected* dan Expectation-Maximization algorithm (algotiirma EM) untuk menduga respon yang tidak lengkap (incomplete data).
 4. AMMI masih berbasis pada respon tunggal seperti tingkat daya hasil. Padahal pada kenyataannya tingkat adaptasi tanaman tidak hanya cukup dilihat dari daya hasil saja melainkan juga harus memperhatikan perkembangan morfologi tanaman maupun daya resistensi tanaman terhadap serangan hama dan penyakit. Dengan demikian respon yang dihadapi dalam kasus ini adalah percobaan lokasi ganda dengan respon ganda. Sehingga diperlukan pendekatan lain yang mampu menarik kesimpulan secara komprehensif dari berbagai respon yang diamati. Beberapa metode telah dipelajari dalam upaya penggabungan respon antara lain menggunakan (i) informasi nilai minimum dan maksimum dari data respon peubah asal (Lawrence et al, 2003; Gani & Duncan, 2004; Kundu, 2004) atau (ii) first principal component (komponen utama pertama) dan jarak Hotelling (Sumertajaya 2005).
 5. Disamping penggabungan respon, ada pula keinginan untuk mengkaji secara lebih mendalam struktur interaksi genotipe \times lingkungan dengan memperhatikan faktor-faktor lingkungan, genotipik dan juga proses fisiologis pertumbuhan dan perkembangan benih. Model AMMI memiliki keterbatasan dalam hal ketidakmampuan menjelaskan pengaruh dari kovariat genotipik dan lingkungan terhadap nyataanya efek interaksi genotipe \times lingkungan. Upaya ini menggabungkan metode AMMI dengan *Structural Equation Modeling* (SEM), yang dikenal dengan nama SEM-AMMI.

6. Kenyataan bahwa pada model anova, terdapat kemungkinan pendugaan nilai komponen ragam baik ragam genetik maupun ragam lingkungan dimungkinkan diperoleh angka yang negatif. Dimana hal ini akan menyulitkan dalam analisis dan interpretasi. Inferensia Bayes memiliki metode-metode yang dapat menghindari permasalahan di atas. Pengembangan model AMMI dengan pendekatan bayes akan dilakukan melalui kajian bayes pada analisis stabilitas genotipe. Sementara ini telah diperoleh informasi awal bahwa pendekatan bayes dengan prior yang tepat dapat meningkatkan keakuratan pendugaan parameter.
7. Masalah lain adalah munculnya pengamatan pencilan. Sebagaimana model multiplikatif berbasis Least Square dan SVD yang lain (diantaranya Analisis Komponen Utama dan Analisis Faktor), AMMI dan Biplot AMMI juga rentan terhadap adanya pencilan. Sedangkan dalam upaya mendapatkan varietas yang unggul, pencilan justru menjadi sesuatu yang berharga, karenanya mengabaikannya tidaklah bijaksana. Untuk itu diperlukan metode yang relatif kekar terhadap adanya pencilan. C. Croux et al (2003) memperkenalkan penggunaan algoritma alternating regression pada model faktor analisis dengan pendekatan robust factorization matrix. Pendekatan ini akan dicoba untuk membangun model robust-AMMI.
8. Beberapa pengamatan tingkat adaptasi adakalanya dicatat secara kualitatif atau pun cacahan. Perluasan dari G-AMMI diperlukan untuk menangani data cacahan, sebaran Poisson menemui fenomena menarik dengan kasus kasus seperti Zero Inflated Poisson ataupun Negative Binomial Counts. Hal ini membawa kita pada suatu kelas pemodelan statistika yang lebih luar yaitu interaksi/asosiasi tabel dua arah baris-kolom Goodman (Goodman Row-Column Two Ways Tabel Assosiasi) yang bersisikan data cacahan. Strategi baru dapat ditempuh dengan memodelkannya melalui pendekatan Reduce-Rank Regression (Yee & Hastie 2003, Yee & Hadi 2011).
9. Integrasi Komputasi metode yang telah Dikembangkan. Ini merupakan kegiatan pengembangan program aplikasi ini ditujukan untuk mengintegrasikan beberapa metode analisis interaksi genotipe interaksi secara komprehensif. Selain itu, kegiatan ini juga dimaksudkan untuk menyediakan software alternatif dengan beberapa keunggulan antara lain (i) dikembangkan menggunakan software R yang gratis, (ii) dilengkapi *Graphical User Interface* (GUI) untuk memudahkan penggunaan, (iii) memiliki kekhususan pada pengembangan metode AMMI.

Model AMMI Kini dan Yang Akan Datang



Gambar 1. Roadmap Penelitian

3

ANALISIS STABILITAS AMMI: DARI INDEKS STABILITAS AMMI KE RESAMPLING BOOTSTRAP

3.1 Indeks Stabilitas AMMI

Indeks stabilitas diperlukan untuk mempermudah melihat tingkat stabilitas suatu genotipe terhadap lingkungan. Sejauh ini model AMMI telah mampu menerangkan pengaruh interaksi genotipe \times lingkungan. Pengembangan juga dilakukan dengan memformulasi index stabilitas dari biplot AMMI. Indeks dibangun berdasarkan konsep jarak, sehingga semakin besar indeks suatu genotipe, maka semakin jauh jarak genotipe dari pusat sumbu koordinat, artinya semakin tidak stabil genotipe tersebut. Jaya (2008), menggunakan indeks stabilitas genotipe ditentukan oleh skor KUI yang dihasilkan dari model AMMI2, sehingga hanya menggunakan skor KUI1 dan skor KUI2. Indeks yang didasarkan

pada dua nilai KUI terbesar tersebut baik digunakan jika persentase keragaman genotipe dan lingkungan yang dapat dijelaskan oleh model AMMI2 besar. Tetapi kurang efektif jika persen keragaman biplot AMMI2 kecil.

Sa'diyah dan Mattjik 2008 mengusulkan indeks yang tidak tergantung pada tingkat keakuratan model AMMI2, yaitu dengan melibatkan semua nilai KUI yang nyata. Dengan tetap mengacu pada konsep jarak index ukur menggunakan jarak euclid titik Genotipe terhadap titik pusat (0.0) pada dimensi penuh yang meliatkan seluruh skor Komponen Utama Interaksi (KUI), dan bukan pada dua dimensi Biplot. Tentu saja masing-masing KUI, katakanlah terdapat m buah KUI yang nyata, tersebut tidaklah setara dalam menentukan besarnya index. Sehingga dalam hal ini digunakan pembobot dari nilai singular dari masing-masing KUI.

3.2 Uji Lanjut Pengaruh Interaksi pada Model AMMI

Namun demikian, model AMMI belum dapat digunakan sebagai penguji lanjut (post-hoc) untuk menguji kontribusi yang diberikan oleh genotipe dan lingkungan terhadap pengaruh interaksi tersebut. Menurut Marasinghe (1980), dalam percobaan faktorial dengan klasifikasi dua faktor (baris dan kolom) jika pengaruh interaksi antara kedua faktor tersebut nyata dan bersifat multipikatif maka perlu dilakukan pengujian lebih lanjut yang disebut dengan pengujian subhipotesis guna mengetahui faktor baris mana yang memberikan kontribusi terhadap pengaruh interaksi antar kedua faktor tersebut. Faktor baris yang nyata tidak berkontribusi terhadap pengaruh interaksi multiplikatif tersebut berarti faktor baris tersebut tidak berinteraksi dengan faktor kolom.

Bila suatu genotipe tertentu nyata berkontribusi terhadap interaksi genotipe \times lingkungan maka genotipe tersebut berinteraksi dengan lingkungan atau bisa dikatakan daya adaptasinya di berbagai lingkungan kurang (cenderung tidak stabil). Sebaliknya, jika genotipe tersebut tidak nyata berkontribusi terhadap interaksi genotipe \times lingkungan maka genotipe tersebut tidak berinteraksi dengan lingkungan atau bisa dikatakan daya adaptasinya di berbagai lingkungan cukup tinggi (cenderung stabil). Rusida (2009) mengintroduksi metode Marasinghe pada Model AMMI untuk menguji sumbangan faktor genotipe pada pengaruh interaksi genotipe \times lingkungan.

Pengujian subhipotesis ini membutuhkan informasi nilai kriteria ujinya untuk mengetahui nilai batasan tolak H_0 dan menarik kesimpulan. Karena nilai kriteria uji tersebut belum diketahui maka salah satu cara guna mengetahui nilai tersebut adalah dengan mengaproksimasi. Pengujian sub-hipotesis pada interaksi

genotipe \times lingkungan dilakukan dengan aproksimasi menggunakan resampling data dengan pengembalian.

Pengujian subhipotesis ini dimulai dengan pengelompokan genotipe untuk memperbanyak ulangan agar nilai statistik uji (Λ) yang dihasilkan beragam. Jika pengelompokan genotipe hasil biplot KUI 1 dan KUI 2 menghasilkan kelompok yang beranggotakan minimal dua genotipe maka sebaiknya pengelompokan genotipe sebelum melakukan pengujian subhipotesis, dilakukan berdasarkan hasil biplot KUI 1 dan KUI 2 (kemiripan karakteristik dari interaksi genotipe \times lingkungan) supaya tidak terjadi perbedaan hasil antara pengelompokan genotipe awal (sebelum pengujian subhipotesis) dengan pengelompokan genotipe berdasarkan hasil pengujian subhipotesis (setelah pengujian subhipotesis).

3.3 Pendugaan Kestabilan Genotipe pada Model AMMI menggunakan Resampling Bootstrap.

Biplot pada analisis AMMI merupakan alat yang digunakan untuk menginterpretasi hasil dari Interaksi Genotipe \times Lingkungan (IGL) dalam model AMMI. Biplot AMMI biasanya berupa biplot antara nilai komponen utama pertama dengan rataan respon (biplot AMMI₁). Biplot antara komponen utama kedua dan nilai komponen pertama (biplot AMMI₂) bisa ditambahkan jika komponen utama kedua ini nyata. Biplot AMMI hanyalah suatu analisis eksplorasi dan tidak menyediakan pengujian hipotesis. Permasalahan yang dapat muncul dalam analisis AMMI adalah menentukan genotipe yang stabil berdasarkan biplot AMMI₂. Kecilnya skor KUI₁ dan KUI₂ menyatakan rendahnya keragaman respon pada suatu genotipe, sehingga genotipe tersebut dapat dikatakan stabil. Semakin besar keragaman suatu genotipe, maka semakin tidak stabil genotipe tersebut. Ukuran keragaman pada biplot AMMI₂ merupakan jarak relatif suatu genotipe terhadap titik pusat (0,0), sehingga dapat mengakibatkan perbedaan interpretasi kestabilan berdasarkan gambar biplot.

Beberapa konsep kestabilan genotipe secara inferensia dalam pemuliaan tanaman telah diperkenalkan, diantaranya adalah pengujian subhipotesis yang digunakan untuk menguji sumbangan faktor genotipe pada pengaruh IGL. Suatu genotipe yang berkontribusi nyata terhadap IGL dikatakan sebagai genotipe yang cenderung tidak stabil. Sebaliknya genotipe yang tidak nyata berkontribusi terhadap IGL dikatakan sebagai genotipe yang cenderung stabil. Pengujian subhipotesis memerlukan nilai statistik uji kemungkinan maksimum dan nilai kriteria uji untuk menarik kesimpulan hasil pengujian. Pengujian subhipotesis

AMMI mampu menguji pengaruh genotipe terhadap IGL dan menghasilkan kelompok genotipe berdasarkan stabilitasnya (Yuliyanti 2009). Pengujian subhipotesis AMMI dinilai kurang efisien dalam menentukan nilai statistik uji kemungkinan maksimum dan sensitif terhadap hasil pengelompokan.

Analisis inferensia lainnya telah dikemukakan oleh Lavoranti (2003) dalam Lavoranti, *et al.* (2007) dengan menggunakan metode resampling bootstrap untuk menguraikan matriks IGL dalam model AMMI. Metode resampling bootstrap AMMI menghasilkan daerah kepercayaan untuk skor KUI_1 dan KUI_2 serta indeks kestabilan masing-masing genotipe dan lingkungan. Dengan adanya daerah kepercayaan dan indeks kestabilan dapat mengurangi ketidakpastian identifikasi dan interpretasi pada saat skor komponen utama bernilai rendah, sehingga metode ini lebih akurat dan terpercaya untuk pemilih varietas dalam program pengembangan genetika.

Pendekatan metode resampling bootstrap yang digunakan oleh Lavoranti untuk model AMMI hanya berdasarkan pada pengaruh interaksi, sehingga selang kepercayaan dan indeks kestabilan yang diperoleh tidak mempertimbangkan pengaruh aditif model. Selain itu, konsep resampling yang digunakan juga akan mempengaruhi struktur genotipe dan lingkungan pada saat menghasilkan skor komponen utama interaksi. Skor komponen utama interaksi merupakan hasil analisis komponen utama matriks IGL yang merupakan matriks sisaan pengaruh aditifnya, sehingga akan lebih akurat jika pengaruh aditif juga dipertimbangkan untuk pendugaan daerah kepercayaan dan kestabilan suatu genotipe.

Novianti (2010) menggunakan metode resampling bootstrap untuk menentukan daerah kepercayaan dan kestabilan pada model AMMI dengan melakukan proses resampling pada data respon. Berdasarkan penduga bootstrap KUI_1^* dan KUI_2^* diketahui bahwa r^* berdistribusi normal lingkaran (circular normal distribution), sehingga diperoleh batas atas dan batas bawah dari r^* . Batas atas r_{atas}^* dan batas bawah r_{bawah}^* ini digunakan sebagai jari-jari lingkaran luar dan lingkaran dalam daerah kepercayaan KUI_1 dan KUI_2 . Genotipe dengan jari-jari lingkaran dalam r_{bawah}^* mendekati titik pusat (0,0) dikategorikan sebagai genotipe yang stabil. Selain berdasarkan jari-jari lingkaran dalam, kestabilan suatu genotipe dengan menggunakan resampling bootstrap juga dipengaruhi oleh simpangan baku r^* . Semakin besar simpangan baku r^* akan memperbesar daerah antara lingkaran luar dan lingkaran dalam yang mengindikasikan semakin tidak stabilnya genotipe tersebut. Analisis AMMI menggunakan metode resampling bootstrap dapat digunakan untuk penelitian yang skor KUI_1 dan KUI_2 genotipenya kecil, sehingga penentuan suatu genotipe yang stabil akan lebih akurat.

Daerah kepercayaan rata-rata skor KUI_1^* dan KUI_2^* untuk metode resampling bootstrap dipengaruhi dan sangat sensitif terhadap posisi titik skor KUI_1 dan KUI_2 , sehingga jarak posisi skor KUI_1^* dan KUI_2^* akan lebih baik digunakan untuk menentukan kestabilan genotipe dibandingkan dengan menggunakan daerah kepercayaan skor KUI_1^* dan KUI_2^* . Analisis AMMI menggunakan resampling bootstrap cocok untuk mendeteksi kestabilan pada percobaan multilokasi apabila ragam genotipe pada biplot kecil.

Berikut ini disajikan beberapa teladan penerapan atas hasil pengembangan analisis stabilitas dengan model AMMI. Artikel-artikel berikut ini adalah hasil penelitian tentang pengembangan analisis stabilitas genotipe menggunakan model AMMI, mulai dari (1) Indeks Stabilitas AMMI untuk Penentuan Stabilitas Genotipe pada Percobaan Multilokasi, (2) Pengembangan Uji Statistika untuk identifikasi genotipe yang memberikan kontribusi terhadap interaksi genotipe x lingkungan pada model AMMI dan (3) Pengembangan metode Pendugaan Daerah Kepercayaan dan Kestabilan Genotipe Pada Model AMMI Menggunakan Metode Resampling Bootstrap.

Halaman ini sengaja dikosongkan

Indeks Stabilitas AMMI untuk Penentuan Stabilitas Genotipe pada Percobaan Multilokasi

Halimatus Sa'diyah¹ dan A. A. Mattjik²

¹Dosen Fakultas Pertanian Universitas Jember

²Guru Besar Departemen Statistika Institut Pertanian Bogor

PENDAHULUAN

Berbagai penelitian perlu dilakukan untuk mendapatkan varietas-varietas unggul, antara lain rekayasa genetika, persilangan antar galur, serta uji daya adaptasi berbagai genotipe pada berbagai kondisi lingkungan (percobaan multilokasi).

Percobaan multilokasi merupakan percobaan yang sering digunakan dalam penelitian pemuliaan tanaman, melibatkan dua faktor utama yaitu genotipe tanaman dan kondisi lingkungan (lokasi: tempat (*site*), musim, perlakuan agronomis (Mattjik, 2005). Dari percobaan ini diharapkan mampu memilah pengaruh utama (genotipe dan lokasi) dan pengaruh interaksi antara genotipe \times lokasi (IGL). Kajian ini penting dalam pemuliaan tanaman karena hasilnya dapat digunakan untuk menduga dan menyeleksi genotipe-genotipe yang berpenampilan stabil (*stability of genotypes*) pada berbagai lingkungan berbeda atau beradaptasi pada suatu lingkungan spesifik (*adaptation of genotypes to specific environmental*). Untuk mampu memilah kedua pengaruh ini dengan baik dibutuhkan pendekatan analisis yang tepat.

Beberapa analisis stabilitas dan adaptabilitas telah berkembang, dengan pendekatan yang berbeda-beda. Analisis AMMI (*Additive Main Effects and Multiplicative Interaction*) merupakan suatu metode multivariat yang mutakhir dalam penelitian pemuliaan tanaman untuk mengkaji IGL pada suatu percobaan multilokasi. Model ini sebenarnya telah dikembangkan oleh Mandel pada tahun 1961 (Husein, *et al.*, 2000). Fleksibilitas dan kemampuan mendekomposisi keragaman pengaruh interaksi model AMMI lebih baik dari model regresi karena analisisnya yang didasarkan pada Analisis Komponen Utama secara statistik telah mampu memilah-milah pengaruh interaksi atas komponen-komponen utama yang bersifat ortogonal. Analisis dilakukan secara serempak dengan menggunakan genotipe dan lingkungan sekaligus, dan bukan secara parsial antara genotipe pada setiap lingkungan dengan indeks lingkungan seperti pada model regresi. Analisis secara serempak seperti ini memungkinkan pengaruh multiplikatif suatu lingkungan turut terdeteksi dan dapat diinterpretasikan, sedangkan penggunaan indeks lingkungan gagal mengkaji pengaruh ini.

ANALISIS STABILITAS

Percobaan multilokasi dilakukan dengan melibatkan berbagai genotipe pada berbagai kondisi lingkungan. Percobaan ini dilakukan seperti halnya rancangan percobaan biasa, hanya saja blok disarangkan ke dalam lingkungan. Data dari percobaan ini memiliki tiga tujuan: a) menduga secara akurat dan memprediksi hasil berdasarkan data percobaan yang terbatas; b) menentukan stabilitas hasil dan pola respon genotipe terhadap lingkungan; c) menyediakan panduan yang *reliable* dalam memilih genotipe atau perlakuan agronomi terbaik untuk penanaman di masa yang akan datang dan di lokasi yang baru (Crossa, *et al.*, 1990). Analisisnya sering disebut analisis ragam gabungan (*Composite Analysis of Variance*). Tentang analisis ragam gabungan dapat dirujuk pada Crossa (1990).

Untuk melakukan analisis stabilitas dari hasil percobaan multilokasi, dihadapi beberapa permasalahan yaitu tentang pengertian stabilitas itu sendiri, serta analisis yang digunakan untuk menentukan stabilitas, yaitu bagaimana menguraikan pengaruh interaksi genotipe dengan lingkungan secara efektif.

Beberapa prosedur statistik telah dikembangkan hingga saat ini untuk menganalisis IGL dan terutama stabilitas hasil pada berbagai kondisi lingkungan, menggunakan beberapa pendekatan yang berbeda. Lin, *et al.* (1986) menggolongkan tiga konsep stabilitas sebagai berikut:

Type 1: Genotipe dianggap stabil jika ragam antar lingkungan kecil. Genotipe stabil memiliki performa yang tidak berubah meskipun kondisi lingkungan bervariasi. Konsep ini berguna antara lain untuk karakter kualitas dan karakter ketahanan penyakit. Tipe ini termasuk stabilitas statis menurut Becker dan Léon (1988), dalam Albert (2004). Parameter yang digunakan untuk menilai stabilitas ini adalah koefisien keragaman (CV) yang digunakan oleh Francis dan Kannenberg (1978) untuk tiap genotipe sebagai parameter stabilitas dan ragam genotipe antar lingkungan, tetapi tidak dapat digunakan untuk menilai adaptasi spesifik.

Type 2: Genotipe dianggap stabil jika responnya terhadap lingkungan paralel terhadap rata-rata respon seluruh genotipe yang dilibatkan pada percobaan. Becker dan Léon (1988), dalam Albert (2004) menggolongkan sebagai stabilitas dinamik atau stabilitas agronomis. Genotipe yang stabil tidak menyimpang dari respon umum lingkungan. Koefisien regresi (*bi*) (Finlay dan Wilkinson, 1963) dan ragam dari stabilitas (Shukla, 1972) dapat digunakan untuk mengukur stabilitas tipe 2 ini.

Type 3: Genotipe dianggap stabil jika kuadrat tengah galat (*residual MS*) dari model regresi pada indeks lingkungan kecil. Indeks lingkungan adalah rataan respon dari semua genotipe di tiap lokasi dikurangi rataan umum (*grand mean*) dari semua genotipe di semua lokasi. Tipe ini juga termasuk stabilitas dinamik menurut Becker dan Léon (1988), dalam Albert (2004). Metode untuk

mendeskrripsikan stabilitas tipe 3 adalah metode Eberhart dan Russell (1966), serta Perkins dan Jinks (1968).

Menurut Westcott (1986), pendekatan regresi adalah metode yang paling berguna untuk bidang pemuliaan tanaman, tetapi metode regresi memiliki keterbatasan biologis dan statistik. Keterbatasan pertama adalah rata-rata genotipe (variabel x) tidak bebas terhadap rata-rata marginal lingkungan (variabel y). Meregresikan satu set variabel terhadap variabel lain yang tidak saling bebas melanggar asumsi regresi (Freeman, 1973). Masalah ini teratasi jika genotipe yang digunakan banyak (15-20 genotipe). Keterbatasan kedua secara statistik adalah bahwa *error* yang berhubungan dengan *slope* genotipe secara statistik tidak saling bebas, karena jumlah kuadrat deviasi dengan derajat bebas $(G-1)$ $(E-2)$, tidak dapat dibagi secara orthogonal antar G buah genotipe (Crossa, 1990). Masalah ketiga adalah penggunaan asumsi bahwa interaksi dan rata-rata lingkungan berhubungan linier. Jika asumsi ini dilanggar, maka efektifitas analisis berkurang dan menghasilkan kesalahan kesimpulan (Westcott, 1986). Hal ini terjadi karena pendekatan ini hanya menjelaskan komponen linier dari pengaruh interaksi sehingga apabila pola interaksi genotipe terhadap lingkungan tidak linier akan menyisakan keragaman yang cukup besar.

Secara umum, berbagai metode statistik yang telah dikemukakan di atas ternyata masih memiliki tingkat keberulangan (*repeatability*) yang relatif rendah, dan memberikan hasil interpretasi yang tidak sama sehingga membingungkan pemulia tanaman dalam menyeleksi genotipe yang stabil atau memiliki adaptasi pada lingkungan spesifik (Lin, *et al.*, 1986).

Kelamahan ini menjadi pemicu berkembangnya metode AMMI (*Additive Main Effects and Multiplicative Interaction*). Tokoh-tokoh yang mengembangkan metode ini antara lain Gauch dan Zobel (1988), serta Zobel, *et al.* (1988), dalam Sumertajaya (1998). AMMI sangat efektif menjelaskan interaksi genotipe dengan lingkungan. Penguraian pengaruh interaksi dilakukan dengan model bilinear, sehingga kesesuaian tempat tumbuh bagi genotipe akan dapat dipetakan dengan jelas. Hal ini terlihat dengan jelas dalam pemetaan genotipe dan lingkungan secara simultan dengan menggunakan BIPLLOT.

ANALISIS AMMI

Analisis AMMI adalah suatu teknik analisis data percobaan dua faktor perlakuan dengan pengaruh utama perlakuan bersifat aditif sedangkan pengaruh interaksi dimodelkan dengan model bilinear. Model AMMI dapat digunakan untuk menganalisis percobaan multilokasi. Pada dasarnya analisis AMMI menggabungkan analisis ragam aditif bagi pengaruh utama perlakuan dengan analisis komponen utama ganda dengan pemodelan bilinear bagi pengaruh interaksi (Gauch, 2006; Crossa, *et al.*, 1990)

Ada tiga manfaat utama penggunaan analisis AMMI, yaitu pertama analisis AMMI dapat digunakan sebagai analisis pendahuluan untuk mencari model yang lebih tepat. Jika tidak ada satupun komponen yang nyata maka pemodelan cukup dengan pengaruh aditif saja. Sebaliknya jika hanya pengaruh ganda saja yang nyata

maka pemodelan sepenuhnya ganda, berarti analisis yang tepat adalah analisis komponen utama saja. Sedangkan jika semua komponen interaksi nyata berarti pengaruh interaksi benar-benar sangat kompleks, tidak memungkinkan dilakukannya pereduksian tanpa kehilangan informasi penting (Bradu dan Bariel, 1978, Gauch, 1985, dalam [Sumertajaya 1998](#)).

Tujuan kedua dari analisis AMMI adalah untuk menjelaskan interaksi galur x lingkungan. AMMI dengan biplotnya meringkas pola hubungan antar galur, antar lingkungan, dan antara galur dan lingkungan (Kempton, 1984, Zobel, *et al.*, 1988, dalam [Sumertajaya 1998](#)).

Kegunaan ketiga adalah meningkatkan akurasi dugaan respon interaksi galur x lingkungan. Hal ini terlaksana jika hanya sedikit komponen AMMI saja yang nyata dan tidak mencakup seluruh jumlah kuadrat interaksi. Dengan sedikitnya komponen yang nyata sama artinya dengan menyatakan bahwa jumlah kuadrat sisanya hanya galat (*noise*) saja. Dengan menghilangkan galat ini berarti dugaan respon dari interaksi galur x lokasi lebih akurat (Zobel, *et al.*, 1988, dalam [Sumertajaya 1998](#)).

Pemodelan bilinear bagi pengaruh interaksi genotipe dengan lokasi pada analisis AMMI dapat dirujuk pada Hadi dan Sa'diyah (2004).

Biplot AMMI

Alat yang digunakan untuk menginterpretasi hasil dari metode AMMI adalah biplot AMMI. Pada dasarnya metode ini berupaya untuk memberikan peragaan grafik terhadap suatu matriks dalam suatu plot dengan menumpangtindihkan vektor-vektor dalam ruang berdimensi dua. Vektor-vektor yang dimaksud yaitu vektor yang mewakili nilai skor komponen lingkungan dan skor komponen genotipe.

Biplot adalah plot antara satu kolom G dengan kolom G yang lain yang ditampilkan secara bersama-sama dengan plot kolom H dengan kolom H yang lain yang bersesuaian dengan kolom G yang diplot ([Jolliffe, 1986](#)). Biplot AMMI2 adalah plot antara skor Komponen utama interaksi terbesar pertama (KUI1) dengan skor Komponen utama interaksi terbesar kedua (KUI2) dari hasil peguraian singular (SVD) matriks interaksi (I). Skor KUI_i untuk genotipe ke-g adalah $u_{gi} \sqrt{\lambda_i}$ dan Skor KUI_i untuk lingkungan ke-l adalah $v_{il} \sqrt{\lambda_i}$ (g=1,...,a; l=1,2,...,b).

Biplot AMMI2 menggambarkan pengaruh interaksi antara genotipe dan lingkungan. Titik-titik amatan yang mempunyai arah yang sama berarti titik-titik amatan tersebut berinteraksi positif (saling menunjang), sedangkan titik-titik yang berbeda arah menunjukkan bahwa titik-titik tersebut berinteraksi negatif.

Memperoleh KUI dari Matriks Interaksi

Misalkan I adalah matriks interaksi dengan ordo a x b dengan a banyak genotipe dan b adalah banyak lingkungan. Dekomposisi matriks interaksi I dengan SVD sebagai berikut

$$I=U\Omega V'$$

Dimana :

U : vektor ciri genotipe berukuran axm

V : vektor ciri lingkungan berukuran mx b

Ω : Matriks diagonal berukuran mxm dengan elemen diagonal adalah nilai singular

λ_i : Nilai singular ke-i

Dari hasil SVD didapatkan skor KUI untuk genotipe dan lingkungan.

Skor KUI Genotipe (G) didapatkan dari: $G=U*\Omega^{1/2}$

Skor KUI Genotipe (G)₁ = $U_1*\lambda_1^{1/2}$

Skor KUI Genotipe (G)₂ = $U_2*\lambda_2^{1/2}$

Skor KUI Lingkungan (L) didapatkan dari $L=V*\Omega^{1/2}$

Skor KUI Lingkungan (L)₁ = $L_1*\lambda_1^{1/2}$

Skor KUI Lingkungan (L)₂ = $L_2*\lambda_2^{1/2}$

Untuk mengetahui genotipe stabil atau tidak, dalam Biplot AMMI diperkenalkan selang kepercayaan elips yang mengadopsi selang kepercayaan uji vektor rata-rata (Sumertajaya, 2005).

Selang Kepercayaan Elips

Dari biplot AMMI2 dapat pula diperoleh gambaran adaptabilitas genotipe (yang berinteraksi khas dengan lokasi tertentu). Makin dekat jarak lokasi dengan genotipe, atau semakin kecil sudut diantara keduanya, maka makin kuat interaksinya. Untuk menguji kestabilan genotipe maka digunakan pendekatan selang kepercayaan sebaran normal ganda yang berbentuk elips pada skor KUI-nya. Semakin stabil suatu genotipe maka titik koordinatnya akan semakin mendekati pusat koordinasi elips. Selang kepercayaan Elips adalah selang kepercayaan pada biplot dengan pusat (0,0) untuk identifikasi genotipe stabil. Jika letak suatu genotype pada biplot berada di luar area elips, maka genotype tersebut dikategorikan sebagai genotype yang tidak stabil, begitu juga sebaliknya. Adapun persamaan untuk mendapatkan jari-jari elips menurut Hadi dan Sa'diyah (2004) adalah sebagai berikut:

$$r_i = \pm \lambda_i \sqrt{\frac{2(k-1)}{n(k-2)} F_{p,n-2,k}}$$

dengan :

- ri : panjang jari-jari,
- i=1 untuk untuk jari-jari panjang,
- i=2 untuk jari-jari pendek

Model AMMI Kini dan Yang Akan Datang

- n : banyaknya pengamatan (genotipe + lingkungan)
 λ_i^2 : akar ciri ke- i dari matriks koragam (S) skor komponen genotipe lingkungan
 λ_i : nilai singular
 $F_{2,n-2;\alpha}$: nilai sebaran F dengan $db_1=2$ dan $db_2=n-2$ pada taraf $\alpha = 5\%$

Indeks Stabilitas AMMI

Indeks stabilitas diperlukan untuk mempermudah melihat tingkat stabilitas suatu genotipe terhadap lingkungan. Indeks dibangun berdasarkan konsep jarak, sehingga semakin besar indeks suatu genotipe, maka semakin jauh jarak genotype dari pusat sumbu koordinat, artinya semakin tidak stabil genotype tersebut. Menurut Jaya (2008), indeks stabilitas genotype ditentukan oleh skor KUI yang dihasilkan dari model AMMI2, sehingga hanya menggunakan skor KUI1 dan skor KUI2. Indeks tersebut didefinisikan sebagai berikut:

$$ISA = \sqrt{\left[\frac{\lambda_1^{1/2}}{\lambda_2^{1/2}} (\text{SkorKUI}_1) \right]^2 + \left[\text{SkorKUI}_2 \right]^2}$$

Indeks yang didasarkan pada dua nilai KUI terbesar tersebut baik digunakan jika persentase keragaman genotype dan lingkungan yang dapat dijelaskan oleh model AMMI2 besar. Tetapi kurang efektif jika persen keragaman biplot AMMI2 kecil. Oleh karena itu, perlu suatu indeks yang tidak tergantung pada tingkat keakuratan model AMMI2, yaitu dengan melibatkan semua nilai KUI yang nyata (misal sebanyak m buah KUI). Mengacu pada konsep jarak, maka untuk menentukan jarak plot Skor Genotipe terhadap titik pusat (0.0) misalkan disebut (J) dengan dapat didefinisikan sebagai berikut:

$$\begin{aligned}
 J &= \sqrt{\lambda_1^{1/2} (\text{SkorKUI}_1)^2 + \dots + \lambda_m^{1/2} (\text{Skor KUI}_m)^2} \\
 &= \sqrt{\frac{\left(\sum_{i=1}^m \lambda_i^{1/2}\right)^2}{\left(\sum_{i=1}^m \lambda_i^{1/2}\right)^2} \left[\lambda_1^{1/2} (\text{SkorKUI}_1)^2 + \dots + \lambda_m^{1/2} (\text{Skor KUI}_m)^2 \right]} \\
 &= \sum_{i=1}^m \lambda_i^{1/2} \sqrt{\left[\frac{\lambda_i^{1/2}}{\sum_{i=1}^m \lambda_i^{1/2}} (\text{SkorKUI}_i) \right]^2 + \left[\frac{\lambda_i^{1/2}}{\sum_{i=1}^m \lambda_i^{1/2}} (\text{SkorKUI}_i) \right]^2}
 \end{aligned}$$

Dimana m adalah banyaknya komponen utama yang nyata. Perhatikan persamaan di atas, nilai $\sum_{i=1}^m \lambda_i^{1/2}$ disebelah kiri dapat dinyatakan sebagai konstanta (C) dan

$$\sqrt{\left[\frac{\lambda_1^{1/2}}{\sum_{i=1}^m \lambda_i^{1/2}} (\text{SkorKUI}_1) \right]^2 + \dots + \left[\frac{\lambda_m^{1/2}}{\sum_{i=1}^m \lambda_i^{1/2}} (\text{SkorKUI}_m) \right]^2}$$

adalah sebuah indeks, yang dapat disebut dengan indeks stabilitas AMMI dari m buah KUI yang nyata ($ISA-m$), sehingga persamaan di atas dapat dituliskan kembali sebagai berikut:

$$J = C * \sqrt{\left[\frac{\lambda_1^{1/2}}{\sum_{i=1}^m \lambda_i^{1/2}} (\text{SkorKUI}_1) \right]^2 + \dots + \left[\frac{\lambda_m^{1/2}}{\sum_{i=1}^m \lambda_i^{1/2}} (\text{SkorKUI}_m) \right]^2}$$

$$= C * ISA$$

Sehingga :

$$ISA = \sqrt{\left[\frac{\lambda_1^{1/2}}{\sum_{i=1}^m \lambda_i^{1/2}} (\text{SkorKUI}_1) \right]^2 + \dots + \left[\frac{\lambda_m^{1/2}}{\sum_{i=1}^m \lambda_i^{1/2}} (\text{SkorKUI}_m) \right]^2}$$

TELADAN PENERAPAN

Data yang digunakan dalam penelitian ini adalah data pemuliaan Jagung Hibrida dari PT. Kreasidharma bekerjasama dengan Bioseed Inc. yang telah dilakukan mulai tanggal 23 Juli 2006 sampai 10 April 2007 pada musim hujan dan kemarau. Percobaan melibatkan 9 genotipe Jagung Hibrida Harapan dan 3 genotipe Jagung Hibrida Komersial. Dalam penelitian ini diambil data pada 16 lingkungan percobaan.

Tabel 1 menunjukkan ada perbedaan rata-rata hasil panen antar genotipe dan rata-rata hasil panen antar lingkungan pada taraf 5%. Selain itu, pengaruh interaksi genotipe \times lingkungan berbeda nyata, sehingga analisis AMMI pada struktur interaksi tersebut akan bermakna. Dari analisis AMMI, penguraian matriks interaksi menghasilkan lima akar ciri yang nyata pada taraf 5% yaitu 0.635, 0.364, 0.279, 0.218 dan 0.158. Kontribusi masing-masing akar ciri terhadap jumlah kuadrat interaksi adalah 32.88%, 18.87%, 14.48%, 11.27% dan 8.20%.

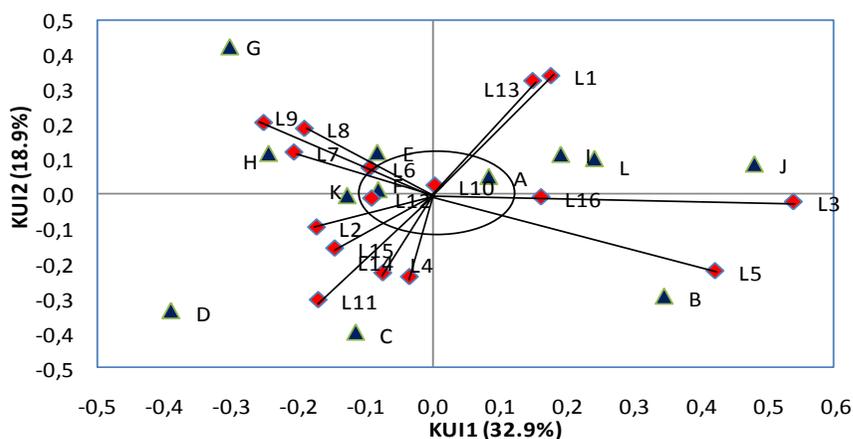
Struktur interaksi antara genotipe dan lingkungan untuk karakter hasil dapat dilihat dari Biplot AMMI-2 pada **Gambar 1** yaitu plot antara KUI_1 dengan KUI_2 .

Model AMMI Kini dan Yang Akan Datang

Hasil biplot ini dapat menggambarkan keragaman interaksi sebesar 51.8%. Persentase keragaman yang dijelaskan relatif besar lebih besar dari 50%. Gambar 1 memperlihatkan bahwa ada dua genotipe yang mempunyai respon relatif stabil terhadap ke-16 lingkungan yaitu genotipe F dan Genotipe A. Genotipe K dan E walaupun berada di luar elips, tetapi jaraknya dari titik pusat tidak terlalu jauh dibandingkan dengan genotipe F dan genotipe A. Jika genotipe-genotipe tersebut berdekatan dengan lingkungan tanam tertentu maka genotipe tersebut dinyatakan spesifik lingkungan menurut karakteristik agronomi yang diamati.

Tabel 1. Hasil Analisis AMMI untuk Karakteristik Hasil

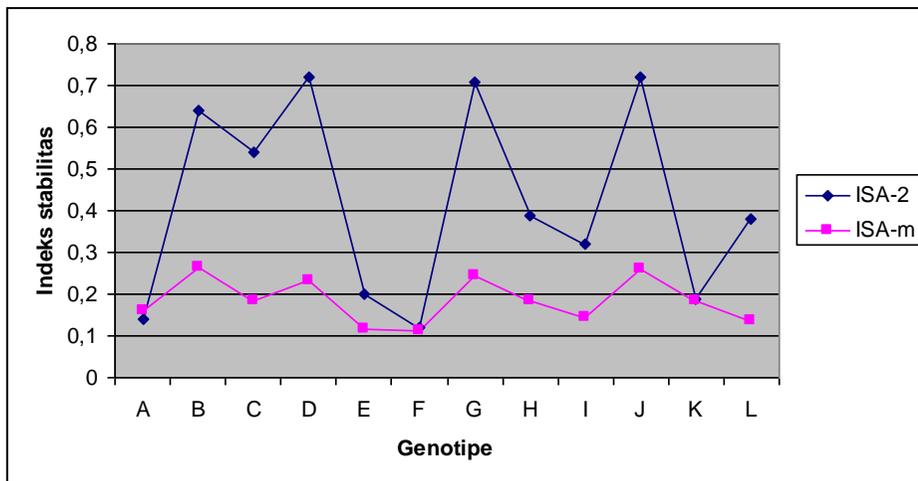
Sumber keragaman	db	JK	KT	F	Nilai-p
Genotipe	11	1.920	0.174	11.1	0.000
Lingkungan	15	66.290	4.420	281.7	0.000
Ulangan(lingkungan)	32	1.790	0.056	3.56	0.000
Interaksi	165	5.790	0.035	2.24	0.000
KUI ₁	25	1.900	0.076	4.85	0.000
KUI ₂	23	1.090	0.048	3.03	0.000
KUI ₃	21	0.840	0.040	2.54	0.000
KUI ₄	19	0.650	0.034	2.19	0.003
KUI ₅	17	0.470	0.028	1.78	0.029
KUI ₆	15	0.310	0.021	1.34	0.177
Sisa	45	0.510	0.011	0.73	0.904
Galat	352	5.520	0.016		
Total	575	81.310	0.141		



Gambar 1. Biplot AMMI 2 Untuk Karakteristik Hasil (51.8%)

Interpretasi Indeks stabilitas AMMI

Indeks stabilitas AMMI menggunakan indeks AMMI-2 dan indeks AMMI-m dapat dilihat pada Gambar 2. Indeks terendah untuk kedua indeks adalah indeks untuk genotype F. Indeks yang rendah mencerminkan bahwa genotype F adalah genotype yang paling stabil pada 16 lingkungan percobaan dibandingkan genotype yang lain. Jika dilihat dari pola dan fluktuasinya, baik indeks ISA_2 maupun ISA-m secara umum memiliki pola yang mirip. Hanya pada indeks untuk genotype E dan genotype L yang menunjukkan perbedaan pola.



Gambar 2. Indeks stabilitas Genotipe

Tabel 3. Indeks Stabilitas AMMI untuk karakter Hasil

Kode Genotipe	Indeks stabilitas AMMI-m		Indeks stabilitas AMMI-2	
	ISA	Rank	ISA	Rank
A	0,159	5	0.140	2
B	0,265	12	0.640	9
C	0,184	6	0.540	8
D	0,232	9	0.720	12
E	0,118	2	0.200	4
F	0,111	1	0.120	1
G	0,244	10	0.710	10
H	0,185	7	0.390	7
I	0,146	4	0.320	5
J	0,261	11	0.720	11
K	0,185	8	0.190	3
L	0,136	3	0.380	6

Dari indeks stabilitas yang dihasilkan, dapat dibuat peringkat atau tingkat stabilitas genotype pada Tabel 3. Hasil peringkat dari kedua Indeks stabilitas

tersebut menghasilkan tingkat stabilitas yang sama untuk genotipe F, G, H dan J. Sedangkan untuk genotipe I, E dan C hanya selisih sedikit. Perbedaan yang tajam adalah pada peringkat untuk genotipe K. Secara keseluruhan, Indeks stabilitas genotipe berdasarkan AMMI-2 maupun AMMI-m menunjukkan beberapa perbedaan, meskipun tidak terlalu besar. Hal ini disebabkan karena indeks AMMI2 hanya menggunakan skor KUI1 dan KUI2, dimana persentase keragaman yang dapat dijelaskan oleh kedua skor KUI tersebut hanya 51.8%, sedangkan Indeks stabilitas menggunakan semua skor KUI yang nyata dan memiliki keakuratan 90% dalam menduga stabilitas yang sesungguhnya.

PENUTUP

AMMI dan grafik biplot AMMI mampu memberikan lebih banyak informasi tentang interaksi genotipe \times lokasi. Namun AMMI tidak lepas dari kelemahan/kerugian. Berikut ini adalah beberapa hal tentang kelemahan/kerugian AMMI dan saran-saran penggunaan AMMI:

1. Data percobaan harus seimbang. Data hilang untuk sementara ini digantikan dengan nilai rata-rata pada lokasi tersebut. Mattjik (1998) merekomendasi penggunaan EM-AMMI untuk menduga data hilang.
2. Total keragaman yang diterangkan mungkin kecil. Sebagai konsekuensi penggunaan biplot KUI, dengan total keragaman yang tergantung akar ciri.
3. Tidak mempunyai ukuran ketidakpastian (measure of uncertainty). Pendekatan AMMI tidak menyediakan pengujian hipotesis seperti pendekatan konvensional lain. Karena itu, AMMI-biplot baik untuk eksplorasi membangkitkan hipotesis daripada pengambilan keputusan hipotesis.
4. Penentuan tingkat stabilitas relatif suatu genotipe yang didasarkan pada Biplot AMMI2, tidak akurat jika total keragaman dari akar ciri pertama dan kedua kecil. Hal ini dapat diatasi dengan menggunakan suatu indeks stabilitas yang melibatkan semua skor komponen yang nyata.

DAFTAR PUSTAKA

- Alberts, M. J. A. 2004. A Comparison of Statistical Methods to Describe Genotype \times Environment Interaction and Yield Stability In Multi-Location Maize Trials. Thesis. Faculty of Agriculture, Department of Plant Sciences, University of the Free State. Bloemfontein.
- Crossa, J. 1990. Statistical Analysis of Multilocation Trials. *Advances In Agronomy*. 44: 55-85.

- Crossa, J., H.G. Gauch, and R.W. Zobel. 1990. Additive Main Effects and Multiplicative Interaction Analysis of Two International Maize Cultivar Trials. *Crop Sci.* 30: 493–500.
- Eberhart, S.A., and W.A. Russel. 1966. Stability Parameters for Comparing Varieties. *Crop Sci.* 6: 36-40
- Francis, T.R., and L.W. Kannenberg. 1978. Yield Stability Studies in Short-Season Maize. I. A. Descriptive Method for Grouping Genotypes. *Can. J. Plant Sci.* 58: 1029-1034.
- Finlay, K.W., and G.N. Wilkinson. 1963. The Analysis of Adaptation in a Plant-Breeding Programme. *Aust. J. Agric. Res.* 14: 742-754.
- Freeman, G.H. 1973. Statistical Methods for The Analysis of Genotype—Environment Interactions. *Heredity*, 31: 339—354.
- Gauch, H.G. 2006. Statistical Analysis of Yield Trials by AMMI and GGE. *Crop Sci.* 46:1488-1500.
- Hadi, A.F., dan H. Sa'diyah. 2004. Model AMMI untuk Analisis Interaksi Genotipa × Lokasi. *Jurnal Ilmu Dasar* vol 5 no. 1: 33-41.
- Hussein, M.A., A. Bjornstad and A.H. Aastveit. 2000. SASG × ESTAB: A SAS Program for Computing Genotype × Environment Stability Statistics. *Agron. J.* 92: 454-459.
- Jaya, I G.N.M. 2008. Analisis Interaksi Genotipe dan Lingkungan Menggunakan Model Persamaan Struktural. Thesis. Sekolah Pascasarjana, IPB. Bogor.
- Jolliffe, I.T. 1986. Principal Component Analysis. Spinger-Verlag, New York.
- Lin, C.S., M.R. Binns, L.P. Lefkovitch. 1986. Stability Analysis: Where Do We Stand?. *Crop Sci.* 26: 894-900.
- Mattjik, A.A. 1998. Aplikasi Analisis Pengaruh Utama Aditif dan Interaksi Ganda (UAIG) pada Data Simulasi. *Forum Statistika Komputasi.* 3: 20-26.
- Mattjik, A.A. 2005. Interaksi Genotipe dan Lingkungan dalam Penyediaan Sumber Daya Unggul. Naskah orasi ilmiah guru besar biometrik Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Pertanian Bogor. Bogor.
- Perkins, J.M., and J.L. Jinks. 1968. Environmental and Genotype-Environmental Components of Variability. III. Multiple Lines and Crosses. *Heredity*, 23: 339-356.
- Shukla, G.K. 1972. Some Statistical Aspects of Partitioning Genotype-Environmental Components of Variability. *Heredity*, 29: 237-245.
- Sumertajaya, I M. 1998. Perbandingan Model AMMI dan Regresi Linier untuk Menerangkan Pengaruh Interaksi Percobaan Multilokasi. Tesis. Jurusan Statistika FMIPA IPB, Bogor
- Sumertajaya I M. 2005. Kajian Pengaruh Interblok dan Interaksi Pada Uji Multilokasi dan Respon Ganda [Disertasi]. Bogor : Sekolah Pascasarjana, Insitut Pertanian Bogor.
- Westcott. B. 1986. Some Methods of Analysing Genotype-Environment Interaction. *Heredity* 56: 243-253.

Halaman ini sengaja dikosongkan

PENDUGAAN DAERAH KEPERCAYAAN DAN KESTABILAN GENOTIPE PADA MODEL AMMI MENGGUNAKAN METODE RESAMPLING BOOTSTRAP

A. A. Mattjik¹⁾, Rusida Yuliyanti²⁾ dan Totong Martono¹⁾

¹⁾Departemen Statistika FMIPA IPB

²⁾Mahasiswa Program Studi Statistika Sekolah Pasca Sarjana IPB

ABSTRACT

Additive Main Effect and Multiplicative Interactions (AMMI) Models have been utilized in analysis of the effects of the Genotype × Environment Interaction (GEI) on multienvironmental trials. However, the models could not be used for post-hoc tests of genotype contribution in the interaction effects. The tests are called tests of subhypothesis. By conducting this test, one could identify the contributing genotypes in GEI. Critical point in this test is unknown, so that it's value must be approximated by resampling method with replacement. This paper used maize data from PT.Kreasidharma and Bioseed Inc. In this data, there are 12 genotypes which are tested in 16 different locations. As the result, the identified contributing genotypes in GEI are BIO 9900, BIO 1169 dan BIO9899.

Keywords : AMMI Model, Tests of Subhypothesis, Critical Point.

PENDAHULUAN

Rancangan percobaan faktorial dua faktor telah banyak diterapkan pada percobaan agronomi yang melibatkan faktor genotipe dan faktor lingkungan, dengan hal menarik yang ingin dikaji adalah pengaruh interaksi kedua faktor tersebut guna mendeteksi genotipe-genotipe yang mempunyai daya adaptasi yang tinggi di berbagai kondisi lingkungan. Model interaksi multiplikatif yang telah banyak digunakan untuk menjelaskan pengaruh interaksi genotipe dengan lingkungan dan juga biasa digunakan untuk analisis kestabilan terhadap hasil percobaan multilokasi adalah model AMMI (*Additive Main Effect and Multiplicative Interactions*). Gauch (1990) mengemukakan bahwa model AMMI merupakan suatu model gabungan dari pengaruh aditif pada analisis ragam dan pengaruh multiplikatif pada analisis komponen utama.

Selama ini model AMMI telah mampu menerangkan pengaruh interaksi genotipe × lingkungan, dan sebenarnya bisa dilakukan pengujian mengenai kontribusi yang diberikan oleh genotipe dan lingkungan terhadap pengaruh interaksi. Menurut Marasinghe (1980), dalam percobaan faktorial dengan

klasifikasi dua faktor (baris dan kolom) jika pengaruh interaksi antara kedua faktor tersebut nyata dan bersifat multipikatif maka perlu dilakukan pengujian lebih lanjut yang disebut dengan pengujian subhipotesis guna mengetahui faktor baris mana yang memberikan kontribusi terhadap pengaruh interaksi antar kedua faktor tersebut. Faktor baris yang nyata tidak berkontribusi terhadap pengaruh interaksi multiplikatif tersebut berarti faktor baris tersebut tidak berinteraksi dengan faktor kolom. Berdasarkan keterangan di atas, apabila genotipe dan lingkungan disetarakan dengan faktor baris dan faktor kolom, dan interaksi genotipe \times lingkungan dimodelkan dengan AMMI, maka kita bisa mengaplikasikan metode Marasinghe untuk menguji sumbangan faktor genotipe pada pengaruh interaksi genotipe \times lingkungan. Bila suatu genotipe tertentu nyata berkontribusi terhadap interaksi genotipe \times lingkungan maka genotipe tersebut berinteraksi dengan lingkungan atau daya adaptasinya di berbagai lingkungan kurang. Sebaliknya, jika genotipe tersebut tidak nyata berkontribusi terhadap interaksi genotipe \times lingkungan maka genotipe tersebut tidak berinteraksi dengan lingkungan atau daya adaptasinya di berbagai lingkungan cukup tinggi.

Pengujian subhipotesis ini membutuhkan informasi nilai kriteria uji untuk mengetahui nilai batasan tolak H_0 dan menarik kesimpulan. Karena nilai kriteria uji tersebut belum diketahui maka salah satu cara guna mengetahui nilai tersebut adalah dengan mengaproksimasi. Proses aproksimasi yang dilakukan menggunakan metode Bootstrap yaitu *me-resampling* data dengan pengembalian.

Tulisan ini bertujuan mendeskripsikan metode pengujian subhipotesis pada model AMMI dan mengimplementasikannya terhadap data percobaan jagung dalam upaya mengidentifikasi genotipe yang berkontribusi terhadap interaksi genotipe \times lingkungan.

METODOLOGI

Menurut Marasinghe (1980), rancangan percobaan faktorial dua faktor dengan interaksi yang terdiri atas a faktor baris dan b faktor kolom, dengan y_{ij} merupakan respon dari faktor baris ke- i pada faktor kolom ke- j , μ adalah nilai rata-rata umum, τ_i adalah pengaruh faktor baris ke- i , β_j adalah pengaruh faktor kolom ke- j , γ_{ij} merupakan pengaruh interaksi antara faktor baris ke- i dan faktor kolom ke- j , dan ε_{ij} adalah pengaruh acak dari faktor baris ke- i pada faktor kolom ke- j yang menyebar Normal $(0, \sigma^2)$, dimodelkan sebagai berikut

$$y_{ij} = \mu + \tau_i + \beta_j + \gamma_{ij} + \varepsilon_{ij} \quad (1)$$

dengan asumsi $\underline{\tau}' \underline{1}_a = 0$, $\underline{\beta}' \underline{1}_b = 0$, $\underline{1}_a' \underline{\Gamma} = \underline{0}'$ dan $\underline{\Gamma} \underline{1}_b = \underline{0}$ jika $\underline{\Gamma} = [\gamma_{ij}]_{a \times b}$.

Marasinghe (1980) mendeskripsikan parameter interaksi γ_{ij} pada model

(1) dalam bentuk bilinear berikut $\gamma_{ij} = \sum_{r=1}^k l_r \alpha_{ir} \theta_{jr}$, dan $k \leq \min(b-1, a-1)$

dengan unsur-unsur dari vektor $\underline{\alpha}_r = \begin{bmatrix} \alpha_{1r} & \alpha_{2r} & \dots & \alpha_{ar} \end{bmatrix}$, $r = 1, 2, \dots, k$, merupakan parameter interaksi faktor baris; sedangkan unsur-unsur vektor $\underline{\theta}_r = \begin{bmatrix} \theta_{1r} & \theta_{2r} & \dots & \theta_{br} \end{bmatrix}$, $r = 1, 2, \dots, k$, merupakan parameter interaksi faktor kolom. Dalam ungkapan bilinear tersebut diasumsikan : $l_1 \geq l_2 \geq \dots \geq l_k$ dan $A'A = B'B = I_k$ dengan $A = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_r & \dots & \alpha_k \end{bmatrix}$ dan $B = \begin{bmatrix} \theta_1 & \theta_2 & \dots & \theta_r & \dots & \theta_k \end{bmatrix}$. Dengan demikian model interaksi multiplikatif dapat ditulis dalam bentuk

$$y_{ij} = \mu + \tau_i + \beta_j + \sum_{r=1}^k l_r \alpha_{ir} \theta_{jr} + \varepsilon_{ij} \quad (2)$$

atau dapat pula ditulis dalam notasi matriks seperti berikut

$$Y = \mu \underline{1}_a \underline{1}_b' + \underline{\tau} \underline{1}_b' + \underline{1}_a \underline{\beta}' + AD \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_k \end{bmatrix}' \underline{B}' + E \quad (3)$$

dengan matriks data Y berordo $a \times b$ dan $D(l_k)$ adalah matriks diagonal berordo k yang unsur-unsur diagonal utamanya ialah l_1, l_2, \dots, l_k , sedangkan E matriks pengaruh acak berordo $a \times b$. Dalam hal ini yang menjadi perhatian pada model (3) ialah pengujian subhipotesis terhadap parameter interaksi pengaruh faktor baris yang dapat diuji dengan menyusun hipotesis berbentuk $H_0 : HA = 0$ dan $H_1 : HA \neq 0$, yang artinya ada sekurang-kurangnya satu ungkapan berbentuk $H_1 : H \underline{\alpha}_r \neq 0$, $1 \leq r \leq k$, dengan H merupakan matriks kontras dan berordo $s \times a$. Hal ini disebabkan karena hipotesis $H_0 : HAD \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_k \end{bmatrix}' = 0$ identik dengan $H_0 : HA = 0$.

Dengan anggapan k sudah ditentukan dan memisalkan $Z = [z_{ij}]$ sebagai matriks interaksi berordo $a \times b$ dengan $z_{ij} = \bar{y}_{ij} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y}_{...}$, maka menurut Marasinghe (1980) hipotesis di atas dapat diuji dengan menggunakan statistik

$$\Lambda = \frac{\sum_{i=1}^a \sum_{j=1}^b z_{ij}^2 - \sum_{r=1}^k \lambda_r}{\sum_{i=1}^a \sum_{j=1}^b z_{ij}^2 - \sum_{r=1}^k \lambda_r^*} \quad (4)$$

dengan λ_r merupakan akar ciri terbesar ke- r dari matriks $Z'Z$ dan λ_r^* merupakan akar ciri terbesar ke- r dari matriks $\begin{bmatrix} -H^{-1}H \\ \underline{Z} \end{bmatrix}'$, sedangkan H^{-1} adalah

matriks kebalikan Moore-Penrose dari matriks H . Hipotesis nol H_0 ditolak jika $\Lambda < q_\phi$ dengan $P_{H_0}(\Lambda < q_\phi) = \phi$. Simulasi Monte Carlo atau Bootstrap dapat digunakan untuk melakukan aproksimasi bagi sebaran uji statistik Λ .

Pada dasarnya analisis AMMI menggabungkan analisis ragam aditif bagi pengaruh utama perlakuan dengan analisis komponen utama dengan pemodelan bilinear bagi pengaruh interaksi yang memanfaatkan penguraian nilai singular (SVD) pada matriks interaksi, sehingga model percobaan faktorial dua faktor menjadi

$$Y = \mu \underline{1}_a \underline{1}_b' + \underline{\tau} \underline{1}_b' + \underline{1}_a \underline{\beta}' + AD \left(\sqrt{\lambda_t} \right) \underline{B}' + E \quad (5)$$

dengan $D \left(\sqrt{\lambda_t} \right)$ adalah matriks diagonal berordo t yang unsur-unsur diagonal utamanya ialah $\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_t}$, $\sqrt{\lambda_t}$ merupakan nilai singular untuk komponen bilinear ke- t (λ_t merupakan akar ciri terbesar ke- t dari matriks ZZ' dan $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_t$), dan δ_{ij} adalah simpangan dari pemodelan bilinear (Crossa dalam Mattjik dan Sumertajaya, 2002). Asumsi-asumsi pada model AMMI identik dengan asumsi pada model interaksi multiplikatif yang diungkapkan oleh Marasinghe (1980) dalam menyusun metode pengujian subhipotesis untuk melakukan identifikasi faktor baris (genotipe dalam model AMMI) yang memberikan kontribusi terhadap interaksi baris \times kolom (genotipe \times lingkungan dalam model AMMI). Oleh karena itu pengujian subhipotesis pada model AMMI dapat dilakukan dengan metode yang diusulkan oleh Marasinghe (1980). Hipotesis nol $H_0 : HA = 0$ lawan $H_1 : HA \neq 0$, dengan H merupakan matriks kontras dan berordo $s \times a$; $A = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_t \end{bmatrix}$ pada model AMMI- t dapat diuji dengan statistik Λ pada persamaan (4) untuk $k = t$ dengan kriteria yang sama dalam menolak hipotesis H_0 .

Bahan

Data yang digunakan dalam penelitian ini merupakan data sekunder hasil percobaan multilokasi tanaman jagung hibrida yang dilakukan oleh Bioseed Genetic International, INC bekerjasama dengan PT. Mitra Kreasidharma pada bulan Juli 2006 sampai dengan April 2007. Percobaan dilakukan di 16 lokasi (sebagai faktor lingkungan) yang tersebar di 6 propinsi dengan 12 genotipe.

Metode Analisis

Pengujian subhipotesis memerlukan nilai statistik uji (Λ) dan nilai kriteria uji (q_ϕ) untuk menarik kesimpulan hasil pengujian. Karena distribusi Λ belum diketahui sehingga didekati dengan distribusi empirik guna mengaproksimasi nilai kriteria uji. Nilai kriteria uji bagi sebaran Λ ditentukan dengan metode Bootstrap.

Hasil metode ini akan baik jika $n > 4$ agar proses Bootstrap, *resampling* data dengan pengembalian menghasilkan variasi data yang memadai. Karena itu dilakukan pengelompokan genotipe menggunakan salah satu cara di bawah ini

1. Acuan kemiripan karakteristik interaksi genotipe \times lingkungan : gunakan biplot melalui AMMI-2 terhadap data.
2. Acuan korelasi peubah respon dengan peubah amatan lainnya : gunakan analisis gerombol hierarki dengan jarak Mahalanobis.

Tahapan yang dilakukan dalam pengujian subhipotesis adalah

1. Persiapan data untuk proses Bootstrap.
Selama $n < 4$ lakukan pengelompokan genotipe dengan salah satu metode yang telah dijelaskan di atas.
2. Pemodelan peubah respon dengan model AMMI.
Jika langkah (1) tidak dikerjakan, maka peubah respon pada model AMMI berupa data semula. Dalam hal lainnya peubah respon tersebut berupa hasil pengelompokan pada langkah (1). Misalkan hasilnya berupa model AMMI- t .
3. Menentukan statistik Λ untuk pengujian subhipotesis $H_0 : HA = 0$ yaitu dengan merumuskan matriks kontras H , menghitung matriks Z , t akar ciri terbesar dari matriks $Z'Z$, katakan $\lambda_1, \lambda_2, \dots, \lambda_t$, menghitung t akar ciri terbesar dari matriks $(-H^{-1}H'ZZ')$, katakan $\lambda_1^*, \lambda_2^*, \dots, \lambda_t^*$.
4. Penentuan nilai kriteria uji q_ϕ dengan metode Bootstrap.

Proses *resampling* dilakukan dengan pengembalian diulang sebanyak 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000 dan 10000 kali untuk melihat kekonsistenan q_ϕ , kemudian menentukan nilai q_ϕ yang konsisten untuk setiap pengujian, berdasarkan hasil dari proses *resampling* tersebut dengan $\phi = 5\%$.

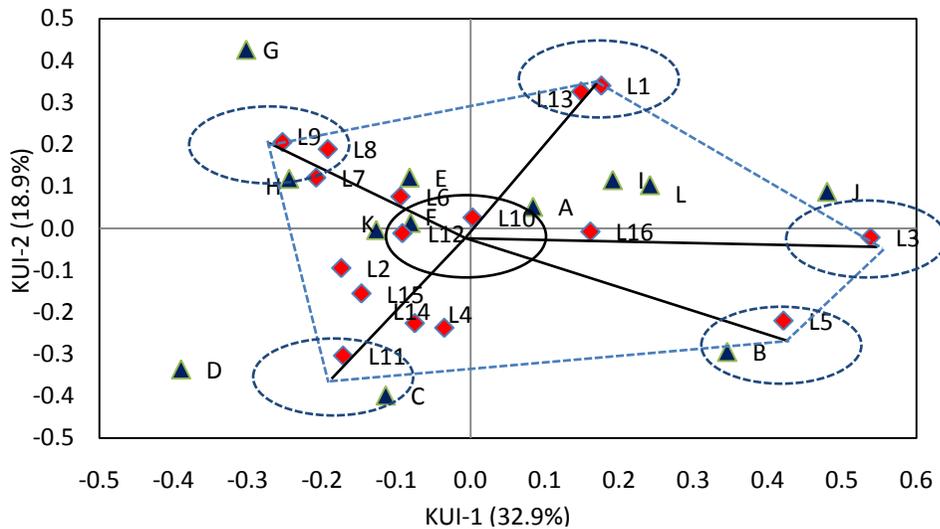
HASIL DAN PEMBAHASAN

Pengelompokan Genotipe

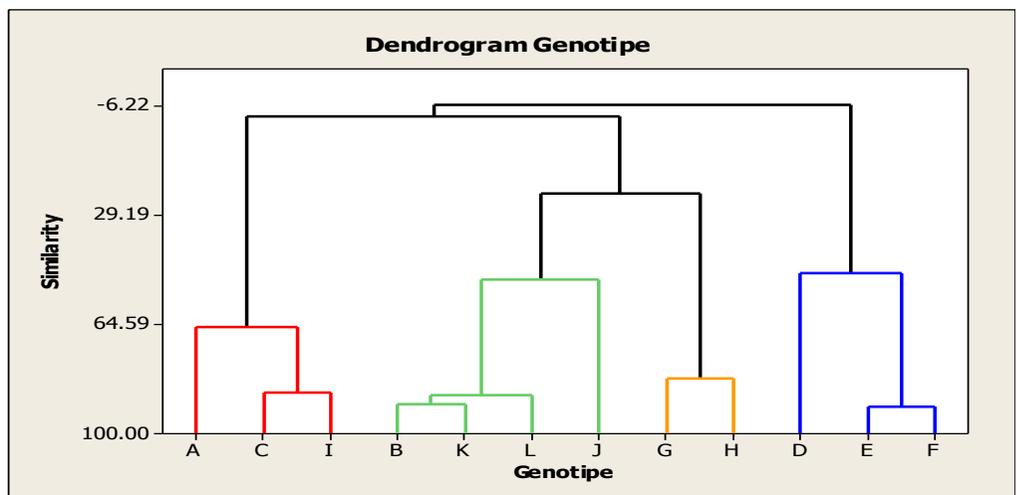
Data percobaan yang digunakan dalam tulisan ini mempunyai ulangan hanya tiga, oleh karena itu dilakukan pengelompokan genotipe guna memperbanyak ulangan agar variasi data yang tersedia dalam proses *resampling* memadai. Pengelompokan genotipe dilakukan dengan dua cara yaitu

1. Pengelompokan genotipe menggunakan biplot melalui AMMI-2. Berdasarkan biplot pada Gambar 1, kedua belas genotipe dikelompokkan menjadi lima kelompok yaitu kelompok I : genotipe A (BIO 9900), F (BC 41399), K (P-12); kelompok II : genotipe B (BIO 1263); kelompok III : genotipe C (BIO 1169) dan D (BC 42521); kelompok IV : genotipe E (BC 42683), G (BC 2630), H (C 42882-A); dan kelompok V : genotipe I (BIO 9899), J (BISI-2), L (C-7).

Model AMMI Kini dan Yang Akan Datang



Gambar 1 Biplot melalui AMMI 2 untuk Hasil Panen



Gambar 2 Dendrogram Genotype

- Pengelompokan genotype menggunakan analisis gerombol hierarki dengan jarak Mahalanobis, dengan acuan korelasi peubah respon hasil panen dengan peubah amatan lainnya, dalam hal ini adalah komponen hasil panen. Menurut Nur et al, (2007), komponen hasil panen yang dapat dijadikan indikator hasil panen adalah jumlah tanaman panen, umur berbunga betina, berat tongkol, dan kadar air panen. Pengelompokan genotype dilakukan menggunakan kriteria kemiripan kadar air saat panen, umur berbunga betina, berat tongkol dan hasil panen, karena variable komponen hasil panen tersebut yang nyata berkorelasi dengan hasil. Berdasarkan Gambar 2, genotype diklasifikasikan ke dalam empat kelompok yaitu kelompok 1 :

genotipe A (BIO 9900), C (BIO 1169), I (BIO9899); kelompok 2 : genotipe B (BIO 1263), J (BISI-2), K (P-12), L (C-7); kelompok 3 : genotipe G (BC 2630), H (C 42882-A); dan kelompok 4 : genotipe D (BC 42521), E (BC 42683), F (BC 41399).

Pengelompokan genotipe dengan biplot melalui AMMI-2 menghasilkan kelompok yang beranggotakan satu genotipe, berarti masih ada kelompok yang mempunyai ulangan tiga ($n < 4$). Sehingga hasil pengelompokan yang digunakan dalam analisis selanjutnya adalah pengelompokan berdasarkan kriteria kemiripan kadar air saat panen, umur berbunga betina, berat tongkol dan hasil panen.

Pengujian Subhipotesis pada Model AMMI 2

Analisis ragam dari hasil panen pada Tabel 1 menunjukkan bahwa pengaruh interaksi kelompok genotipe x lingkungan nyata pada taraf 5% berarti kelompok genotipe memberikan respon hasil panen yang tidak sama di lingkungan yang berbeda.

Tabel 1 Analisis Ragam dari Hasil Panen

Sumber Keragaman	db	Jumlah Kuadrat	Kuadrat Tengah	F	Nilai-p
Kelompok Genotipe	3	16.402	5.467	15.29	0.000
Lingkungan	15	796.533	53.102	148.49	0.000
Kelompok Genotipe*Lingkungan	45	31.380	0.697	1.95	0.002
Sisaan	128	45.774	0.358		
Total	191	890.088			

Pengaruh interaksi yang nyata pada percobaan ini biasanya dimodelkan dengan pola interaksi multiplikatif yaitu dengan model AMMI untuk mengetahui struktur interaksi kelompok genotipe x lingkungan. Berdasarkan Tabel 2 di bawah ini, penguraian pengaruh interaksi kelompok genotipe x lingkungan untuk respon hasil panen nyata sampai KUI-2 dengan 90.77% keragaman yang mampu dijelaskan sehingga model AMMI yang digunakan cukup sampai AMMI-2 dengan model berikut :

$$\hat{y}_{ij} = \hat{\mu} + \hat{\tau}_i + \hat{\beta}_j + \sum_{r=1}^2 \sqrt{\lambda_r} \alpha_{ir} \theta_{jr} + \delta_{ij} + \varepsilon_{ij} \tag{6}$$

Tabel 2 Persentase Keragaman Interaksi Kelompok Genotipe x Lingkungan

Komponen AMMI	Nilai Singular	Keragaman Interaksi Daya Hasil		Signifikansi
		Persentase	Persentase Kumulatif	
KUI -1	2.5725	61.47	61.47	*
KUI -2	1.7763	29.31	90.77	*
KUI -3	0.9967	9.23	100.00	**

*) nyata pada $\alpha = 5\%$, **) tidak nyata pada $\alpha = 5\%$, KUI : Komponen Utama Interaksi

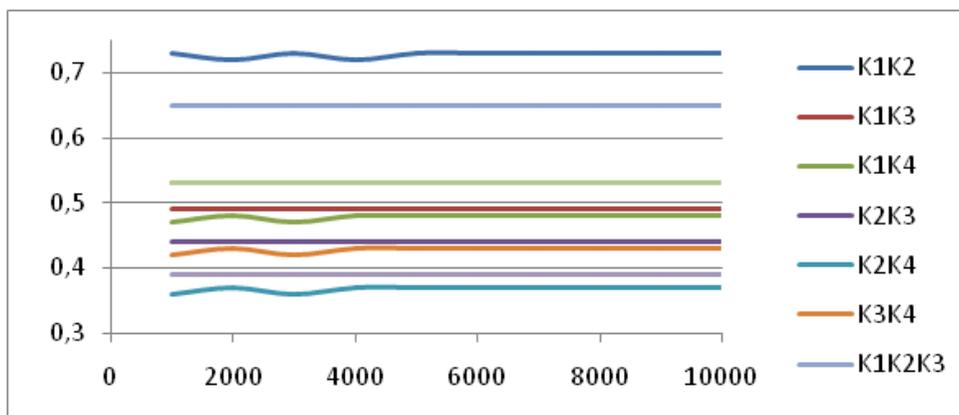
Model AMMI Kini dan Yang Akan Datang

Hasil analisis ragam di atas menyatakan bahwa pengaruh interaksi kelompok genotipe \times lingkungan nyata dan pola interaksinya signifikan pada model AMMI-2. Karena pengaruh interaksi kelompok genotipe \times lingkungan nyata maka untuk mengetahui kelompok genotipe yang berkontribusi terhadap pengaruh interaksi tersebut dilakukan pengujian subhipotesis antar kelompok genotipe, dengan matriks kontras H yang diuji tertera pada Tabel 3.

Tabel 3 Matriks Kontras Pengujian Subhipotesis

Kelompok Genotipe yang Dibandingkan	Matriks Kontras (H)
K1K2	[1 -1 0 0]
K1K3	[1 0 -1 0]
K1K4	[1 0 0 -1]
K2K3	[0 1 -1 0]
K2K4	[0 1 0 -1]
K3K4	[0 0 1 -1]
K1K2K3	[2 -1 -1 0]
K1K2K4	[2 -1 0 -1]
K1K3K4	[2 0 -1 -1]
K2K3K4	[0 2 -1 -1]

Pengujian tersebut memerlukan nilai kriteria uji untuk mengetahui nilai batasan tolak H_0 dan menarik kesimpulan. Karena nilai kriteria uji tersebut belum diketahui maka salah satu cara guna mengetahui nilai tersebut adalah dengan mengaproksimasi. Proses aproksimasi yang dilakukan menggunakan metode Bootstrap yaitu *me-resampling* data dengan pengembalian. Nilai kriteria uji yang dihasilkan pada proses *resampling* yang diulang sebanyak c kali merupakan nilai statistik Λ yang ke- $(5\% \times c)$ dari hasil *resampling* yang telah diurutkan.



Gambar 3 Kekonsistenan Nilai Kriteria Uji

Penentuan nilai kriteria uji adalah dengan memilih nilai kriteria uji hasil aproksimasi yang nilainya telah konsisten pada proses *resampling* yang diulang berapapun. Kekonsistenan nilai kriteria uji terlihat pada Gambar 3, dimana nilai kriteria uji untuk pengujian K1K3, K2K3, K1K2K3, K1K2K4, K1K3K4 dan K2K3K4 telah konsisten pada proses *resampling* yang diulang 1000 kali. Sedangkan nilai kriteria uji pada pengujian K1K2 konsisten pada proses *resampling* yang diulang 5000 kali, untuk pengujian K1K4, K2K4, dan K3K4 nilai kriteria ujinya konsisten pada proses *resampling* yang diulang 4000 kali.

Setelah mendapatkan nilai kriteria uji di setiap pengujian tersebut. pengujian subhipotesis yang dilakukan antar kelompok genotipe telah bisa diambil keputusan apakah terima atau tolak H_0 dengan cara membandingkan nilai statistik uji (Λ) terhadap nilai kriteria uji yang diperoleh dari proses aproksimasi untuk setiap pengujian, seperti tercantum pada Tabel 4.

Tabel 4 Hasil Pengujian Subhipotesis pada AMMI 2

Pengujian	Nilai Λ AMMI 2		Nilai kriteria uji	Keputusan
K1K2	0.70	<	0.73	Tolak H_0
K1K3	0.42	<	0.49	Tolak H_0
K1K4	0.47	<	0.48	Tolak H_0
K2K3	0.72	>	0.44	Terima H_0
K2K4	0.83	>	0.37	Terima H_0
K3K4	0.74	>	0.43	Terima H_0
K1K2K3	0.64	<	0.65	Tolak H_0
K1K2K4	0.51	<	0.53	Tolak H_0
K1K3K4	0.50	<	0.53	Tolak H_0
K2K3K4	0.78	>	0.39	Terima H_0

Pengujian K1K2, K1K3, K1K4, K1K2K3, K1K2K4 dan K1K3K4 mempunyai nilai Λ yang lebih kecil daripada nilai kriteria ujinya sehingga tolak H_0 , sedangkan nilai Λ pengujian K2K3, K2K4, K3K4 dan K2K3K4 lebih besar daripada nilai kriteria ujinya sehingga terima H_0 pada taraf nyata 5%. Berdasarkan hasil pengujian ini berarti :

- a. K1 saja yang nyata berkontribusi terhadap interaksi kelompok genotipe \times lingkungan berarti genotipe A (BIO 9900), C (BIO 1169) dan I (BIO9899) berinteraksi dengan lingkungan sehingga relatif kurang mampu beradaptasi.
- b. K2, K3 dan K4 tidak nyata memberikan kontribusi terhadap interaksi kelompok genotipe \times lingkungan atau hanya menyumbangkan pengaruh aditif (utama) saja sehingga genotipe-genotipe B (BIO 1263), J (BISI-2), K (P-12), L (C-7), G (BC 2630), H (C 42882-A), D (BC 42521), E (BC 42683) dan F (BC 41399) tidak berinteraksi dengan lingkungan dan relatif bisa ditanam dan beradaptasi dengan baik di semua lingkungan yang dicobakan.

KESIMPULAN

Metode subhipotesis Marasinghe dapat digunakan untuk menguji pengaruh genotipe terhadap interaksi genotipe × lingkungan pada model AMMI. Matriks kontras H pada hipotesis nol H_0 merepresentasikan tujuan pengujian tersebut. Pengujian subhipotesis terhadap data percobaan jagung menunjukkan bahwa genotipe yang memberikan kontribusi terhadap interaksi genotipe × lingkungan adalah genotipe A (BIO 9900), C (BIO 1169) dan I (BIO9899).

DAFTAR PUSTAKA

- Gauch, J.R. 1990.** Full and Reduced Models for Yield Trials. *Theoretical and Applied Genetics*. 80: 153-160.
- Marasinghe, M.G. 1980.** Testing Subhypothesis In The Multiplicative Interaction Model [dissertation]. Kansas : Department of Statistics. Kansas State University.
- Mattjik, A.A., dan Sumertajaya, I.M. 2002.** Perancangan Percobaan dengan Aplikasi SAS dan Minitab. Bogor : IPB Press.
- Nur, A., Isaeni, M., Iriany, R.N., dan Takdir, A. 2007.** Stabilitas Komponen Hasil sebagai Indikator Stabilitas Hasil Genotipe Jagung Hibrida. *Penelitian Pertanian Tanaman Pangan* Vol. 26 No. 2 : 106-107.

PENDUGAAN DAERAH KEPERCAYAAN DAN KESTABILAN GENOTIPE PADA MODEL AMMI MENGGUNAKAN METODE RESAMPLING BOOTSTRAP

Pepi Novianti¹⁾, A A Mattjik²⁾ dan I M Sumertajaya²⁾

¹⁾Mahasiswa Program Studi Statistika Sekolah Pasca Sarjana IPB, Dosen Universitas Bengkulu

²⁾Departemen Statistika FMIPA IPB

1. Pendahuluan

Analisis yang banyak digunakan untuk mengkaji interaksi antara genotipe dan lingkungan dalam pemuliaan tanaman adalah analisis *Additive Main Effects Multiplicative Interaction* (AMMI). Pada dasarnya analisis AMMI menggabungkan analisis ragam aditif bagi pengaruh utama perlakuan dan analisis komponen utama ganda dengan pemodelan bilinear bagi pengaruh interaksi (Mattjik & Sumertajaya 2006). Struktur interaksi pada model AMMI diuraikan dari matriks sisaan komponen aditif dengan memanfaatkan sifat matematis penguraian nilai singular (SVD). SVD merupakan pendekatan kuadrat terkecil dengan reduksi dimensi (pangkat matriks) data yang terbaik dan menyediakan penyajian secara grafis yang dikenal secara luas dengan nama Biplot.

Biplot AMMI merupakan alat yang digunakan untuk menginterpretasi hasil dari interaksi genotipe \times lingkungan (IGL) dalam model AMMI. Analisis ini bertujuan memvisualisasikan hubungan antara rata-rata, skor komponen utama interaksi pertama (KUI_1) dan komponen utama interaksi kedua (KUI_2) ke dalam grafik berdimensi dua. Grafik antara rata-rata dan KUI_1 disebut biplot $AMMI_1$, sedangkan grafik antar KUI_1 dan KUI_2 disebut biplot $AMMI_2$. Biplot AMMI hanyalah suatu analisis eksplorasi dan tidak menyediakan pengujian hipotesis. Skor KUI_1 dan KUI_2 yang kecil menyatakan rendahnya keragaman respon pada suatu genotipe, sehingga genotipe tersebut dapat dikatakan stabil. Semakin besar keragaman suatu genotipe, maka semakin tidak stabil genotipe tersebut. Ukuran keragaman pada biplot $AMMI_2$ merupakan jarak relatif suatu genotipe terhadap titik pusat (0,0), sehingga dapat mengakibatkan perbedaan interpretasi dari gambar biplot.

Beberapa konsep kestabilan genotipe secara inferensia dalam pemuliaan tanaman telah diperkenalkan, diantaranya adalah pengujian subhipotesis oleh Yulianti (2009) dan Metode resampling bootstrap oleh Lavoranti *et al.* (2007). Pengujian subhipotesis digunakan untuk menguji sumbangan faktor genotipe pada pengaruh IGL, tetapi Pengujian subhipotesis AMMI dinilai kurang efisien dalam menentukan nilai statistik Λ (uji kemungkinan maksimum) dan sensitif terhadap hasil pengelompokan. Metode resampling bootstrap digunakan untuk menduga daerah kepercayaan dan indeks kestabilan berdasarkan pada pengaruh interaksi, sehingga selang kepercayaan dan yang diperoleh tidak mempertimbangkan pengaruh aditif model. Selain itu, konsep resampling yang digunakan mempengaruhi struktur genotipe dan lingkungan pada saat menghasilkan skor komponen utama interaksi.

Skor komponen utama interaksi merupakan hasil analisis komponen utama matriks IGL yang merupakan matriks sisaan pengaruh aditifnya, sehingga akan lebih akurat jika pengaruh aditif juga dipertimbangkan untuk pendugaan daerah kepercayaan dan kestabilan suatu genotipe. Oleh karena itu akan digunakan metode resampling Bootstrap untuk menentukan daerah kepercayaan dan kestabilan pada model AMMI dengan melakukan proses resampling pada data respon.

2. Tinjauan Pustaka

2.1 Additive Main Effect Multiplication Interaction (AMMI)

Additive Main Effect Multiplicative Interaction (AMMI) adalah suatu teknik analisis data percobaan dua faktor perlakuan dengan pengaruh utama perlakuan bersifat aditif sedangkan pengaruh interaksi dimodelkan dengan model bilinear. Pada dasarnya AMMI menggabungkan analisis ragam aditif bagi pengaruh utama perlakuan dan analisis komponen utama dengan pemodelan bilinear bagi pengaruh interaksi (Mattjik & Sumertajaya 2006).

Model aditif untuk a genotipe dan b lokasi pada model AMMI adalah:

$$Y_{ge} = \mu + \alpha_g + \beta_e + \sum_{k=1}^m \sqrt{\lambda_k} \phi_{gk} \rho_{ek} + \varepsilon_{ge}$$

dimana Y_{ge} adalah respon untuk genotipe ke- g pada lokasi ke- e , $g = 1, 2, \dots, a$, $e = 1, 2, \dots, b$, μ adalah rata-rata umum, α_g adalah pengaruh genotip ke- g , β_e adalah pengaruh lokasi ke- e , dan ε_{ge} adalah galat acak. Syarat untuk sifat

multiplikatif interaksi harus memenuhi kendala $\lambda_1 > \dots > \lambda_m > 0$, $\sum_{k=1}^m \varphi_{gk} = \sum_{k=1}^m \rho_{ek} = 1$ dan $\sum_{g=1}^a \varphi_{gk} \varphi_{gk} = \sum_{e=1}^b \rho_{ek} \rho_{ek} = 0$, $k \neq k$ (Cornelius 1993).

Pengaruh komponen aditif dianalisis menggunakan analisis ragam untuk rancangan percobaan faktorial dengan dua faktor. Apabila interaksi antara GE berpengaruh nyata, maka dikatakan bahwa genotipe dan lingkungan berinteraksi. Interaksi antara genotipe dan lingkungan dapat diketahui dengan menggunakan penguraian nilai singular untuk matriks interaksi Z .

$$Z = \begin{pmatrix} Y_{11} & \dots & Y_{1b} \\ \vdots & \ddots & \vdots \\ Y_{a1} & \dots & Y_{ab} \end{pmatrix}$$

pengaruh interaksi genotipe ke- g pada lokasi ke- e diduga dengan

$$Y_{ge} = \bar{y}_{ge} - \bar{y}_{g.} - \bar{y}_{.e} + \bar{y}_{..}$$

Penguraian matriks Z dilakukan dengan teknik *SVD* dan menggunakan konsep analisis komponen utama (AKU), sehingga analisis model AMMI dapat disajikan dalam grafik berdimensi dua atau disebut dengan biplot AMMI. Penguraian nilai singular untuk matriks pengaruh interaksi Z adalah memodelkan matriks tersebut sebagai berikut:

$$Z = ULA$$

dimana Z adalah matriks interaksi, L adalah matriks diagonal nilai singular dari ZZ , $L = D(\sqrt{\lambda_n})_{m \times m}$, selanjutnya disebut nilai singular. A dan U adalah matriks ortonormal $AA = UU = I_r$. $A = \{a_1, a_2, \dots, a_n\}$ adalah vector ciri ZZ sedangkan U diperoleh dari

$$U = ZAL^{-1} = \left(\frac{Za_1}{\sqrt{\lambda_1}}, \frac{Za_2}{\sqrt{\lambda_2}}, \dots, \frac{Za_n}{\sqrt{\lambda_n}} \right)$$

Dengan mendefinisikan $L = L^\alpha L^{1-\alpha}$ dan $\alpha = 1/2$ diperoleh

$$G = UL^\alpha = \begin{pmatrix} \varphi_{11} & \dots & \varphi_{1n} \\ \vdots & \ddots & \vdots \\ \varphi_{a1} & \dots & \varphi_{an} \end{pmatrix}$$

$$H = AL^{1-\alpha} = \begin{pmatrix} \rho_{11} & \dots & \rho_{b1} \\ \vdots & \ddots & \vdots \\ \rho_{1n} & \dots & \rho_{bn} \end{pmatrix}$$

maka penguraian nilai singular tersebut dapat ditulis:

$$Z = GH$$

dimana kolom-kolom matriks G adalah skor komponen untuk genotipe sedangkan kolom-kolom matriks H adalah skor komponen untuk lingkungan, sehingga interaksi antara GE dapat dinyatakan sebagai:

$$Y_{ge} = \sum_{j=1}^n \sqrt{\lambda_j} \varphi_{gj} \rho_{ej} + \delta_{ge}$$

dimana $\sqrt{\lambda_n}$ nilai singular untuk komponen bilinear ke- n (λ_n adalah akar ciri ZZ), $\lambda_1 \geq \dots \geq \lambda_n$; φ_{gn} adalah pengaruh ganda genotipe ke- g melalui komponen bilinear ke- n ; ρ_{en} adalah pengaruh ganda lingkungan ke- e melalui komponen bilinear ke- n .

Alat yang digunakan untuk menginterpretasi hasil dari model AMMI adalah biplot. Pada dasarnya metode ini adalah upaya untuk memberikan visual grafik dari suatu matriks dalam suatu plot dengan menumpangtindihkan vektor-vektor dalam ruang berdimensi dua. Vektor-vektor yang dimaksud yaitu vektor yang mewakili nilai skor komponen genotipe dan lingkungan. Biplot pada analisis AMMI biasanya berupa biplot pada nilai komponen utama pertama ($KU1_1$) dengan rataan respon (Biplot $AMMI_1$) dan biplot antara komponen utama kedua dan nilai komponen utama pertama (Biplot $AMMI_2$) bisa ditambahkan jika komponen utama kedua ini signifikan (Hadi 2006).

Kestabilan genotipe tanaman dapat diklasifikasikan menjadi genotipe stabil dan genotipe spesifik lingkungan. Genotipe stabil adalah genotipe yang memiliki daya adaptasi tinggi terhadap kondisi lingkungan dengan kata lain genotipe stabil adalah genotipe yang memiliki peringkat sama pada berbagai kondisi lingkungan. Sedangkan genotipe spesifik lingkungan adalah genotipe yang hanya memberikan respon baik terhadap kondisi lingkungan tertentu (Sumertajaya 2005).

2.2 Resampling Bootstrap

Bootstrap merupakan bagian dari metode yang melakukan penarikan contoh berulang dari himpunan data asli sehingga disebut prosedur resampling. Ide pokok resampling bootstrap sederhana, namun dilakukan berulang-ulang sehingga rumit dalam hal perhitungan. Saat ini penerapan metode bootstrap banyak digunakan diberbagai disiplin ilmu dikarenakan semakin berkembangnya teknologi komputer (Chernick 2008).

Definisi formal untuk resampling bootstrap diilustrasikan sebagai berikut: Misalkan suatu vektor acak berdistribusi identik dan saling bebas (IID) dengan n sampel $X = (x_1, x_2, \dots, x_n)$ dan penduga parameternya adalah $\hat{\theta} = s(x)$,

prosedur untuk menduga ketepatan dari $\hat{\theta}$ diperoleh dari fungsi sebaran empiris \hat{F} . Secara empiris sebaran ini menyatakan peluang untuk masing-masing pengamatan dari vektor acak X_i adalah $1/n$, untuk $i = 1, 2, \dots, n$. Contoh bootstrap merupakan pengambilan contoh acak sebanyak n kali dari \hat{F} , yaitu $X^* = (x_1^*, x_2^*, \dots, x_n^*)$,

$$\hat{F} \rightarrow (x_1^*, x_2^*, \dots, x_n^*)$$

X^* bukan suatu data asli, tetapi menyatakan data hasil resampling dari X .

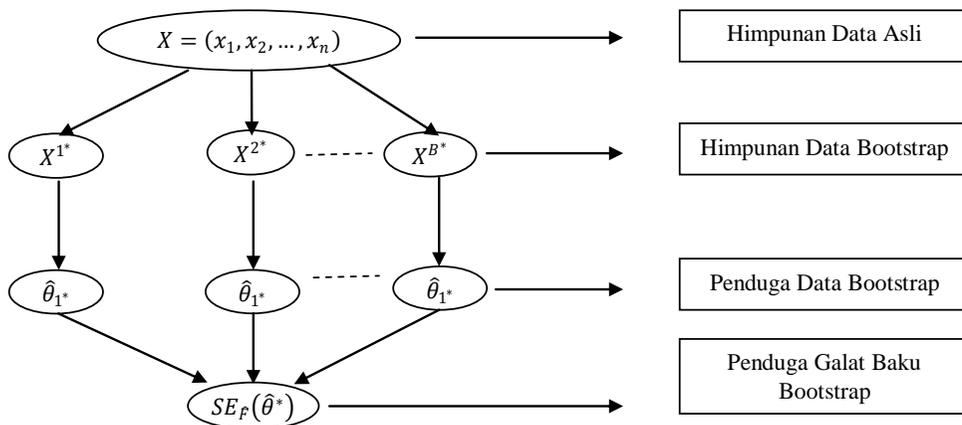
Satu set himpunan data bootstrap memiliki satu nilai dugaan $\hat{\theta}$, yaitu

$$\hat{\theta}^* = s(x^*)$$

Nilai $s(x^*)$ merupakan hasil dari penggunaan fungsi yang sama dari $s(\cdot)$ pada x^* seperti yang digunakan untuk x . Misalkan $s(x)$ merupakan rata-rata contoh $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$, maka $s(x^*)$ adalah rata-rata contoh data bootstrap $\bar{x}^* = \frac{\sum_{i=1}^n x_i^*}{n}$.

Penduga bootstrap $se_F(\hat{\theta})$ merupakan galat baku dari $\hat{\theta}$, yaitu penduga yang menggunakan fungsi sebaran empiris \hat{F} dari distribusi F yang tak diketahui. Penduga bootstrap $se_F(\hat{\theta})$ dinotasikan dengan $se_{\hat{F}}(\hat{\theta}^*)$, yaitu penduga galat baku dari $\hat{\theta}$ untuk himpunan data berukuran n yang diambil secara acak dari sebaran \hat{F} (Efron 1993).

Langkah pendugaan bootstrap dimulai dengan menarik beberapa contoh bootstrap yang saling bebas, menghitung penduga dari ulangan bootstrap dan menduga galat baku dari $\hat{\theta}$ menggunakan galat baku empiris dari ulangan bootstrap. Ilustrasi dari pendugaan galat baku bootstrap disajikan pada Gambar 1.



Gambar 1 Langkah pendugaan galat baku bootstrap $se_{\hat{F}}(\hat{\theta}^*)$

Pendugaan galat baku $se_F(\hat{\theta})$ menggunakan simpangan baku contoh sebanyak B^* ulangan dan dihitung sebagai berikut:

$$\widehat{se}_B = \left\{ \frac{\sum_{b=1}^B [\hat{\theta}^*(b) - \hat{\theta}^*(\cdot)]^2}{B - 1} \right\}^{1/2}$$

dimana

$$\hat{\theta}^*(\cdot) = \frac{\sum_{b=1}^B \hat{\theta}^*(b)}{B}$$

Limit dari \widehat{se}_B dimana B menuju tak hingga disebut penduga bootstrap ideal dari $se_F(\hat{\theta})$,

$$\lim_{B \rightarrow \infty} \widehat{se}_B = se_{\hat{F}} = se_{\hat{F}}(\hat{\theta}^*)$$

Penduga bootstrap ideal $se_{\hat{F}}(\hat{\theta}^*)$ dengan pendekatannya dari \widehat{se}_B biasa juga disebut penduga bootstrap nonparametrik.

Galat baku bootstrap digunakan untuk menyatakan pendekatan selang kepercayaan terhadap parameter θ . Misalkan suatu penduga $\hat{\theta}$ dan penduga galat baku \widehat{se} , maka selang kepercayaan $(100 - \alpha)\%$ untuk θ adalah

$$\hat{\theta} \pm z^{(\alpha)} \widehat{se} \tag{1}$$

dimana $z^{(\alpha)}$ merupakan sebaran normal baku dengan peluang $(100 - \alpha)\%$. persamaan (1) disebut penduga selang atau selang kepercayaan untuk θ .

Bootstrap digunakan bukan untuk menghasilkan satu penduga titik terbaik, tetapi untuk menduga keakuratan penduga parameter. Bootstrap diselesaikan dengan menentukan contoh bootstrap yang digunakan untuk menduga galat baku. Bootstrap tidak membutuhkan rumus analitik yang rumit untuk pendugaan dan dapat digunakan selama masih ada metode komputasi untuk mendapatkan penduga. Hal ini berarti bahwa hanya dibutuhkan penggabungan perhitungan iterasi menggunakan komputer untuk mendapatkan penduga parameter.

3. Hasil dan Pembahasan

Penelitian menggunakan data sekunder yang diperoleh dari hasil penelitian Konsorsium Padi Nasional tahun 2008, yaitu penelitian interaksi antara genotipe dengan lingkungan pada galur harapan padi sawah. Respon yang diukur adalah hasil produksi padi (Ton/Ha) pada 2 set data. Setiap set data terdiri dari 14 galur padi yang ditanam di 20 lingkungan. Adapun desain rancangan yang digunakan

dalam penelitian tersebut adalah Rancangan Acak Kelompok Lengkap dengan 3 ulangan. Kode galur padi dan lingkungan untuk data set 1 disajikan pada tabel berikut:

Tabel 1 Kode galur harapan padi sawah

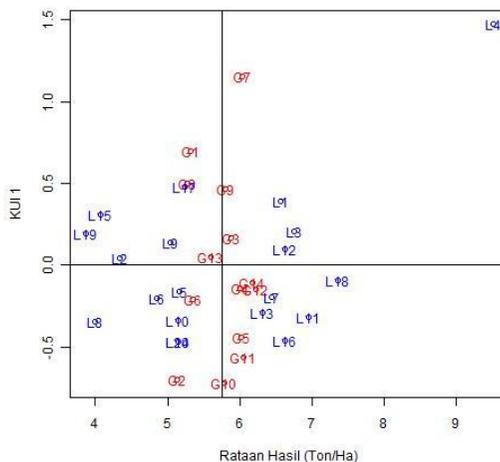
Kode	Genotipe	Kode	Genotipe
G1	IPB-3(IPB102-F-2-1)	G8	IPB-6(IPB107-F-8-3)
G2	BIO-8-AC-BLB/BLASS-05	G9	BP3300-2C-2-3
G3	B10531E-KN-14-3-0-L-R-B376-1	G10	OBS 1739/PSJ
G4	OBS 1735/PSJ	G11	B10531E-KN-14-1-0-L-R-B375-12
G5	BP11252-2-PN-12-2-2-1	G12	CIHERANG
G6	BIO-8-AC-BLB-05	G13	INPARI-1
G7	OBS 1740/PSJ	G14	CIMELATI

Tabel 2 Kode lokasi tanam

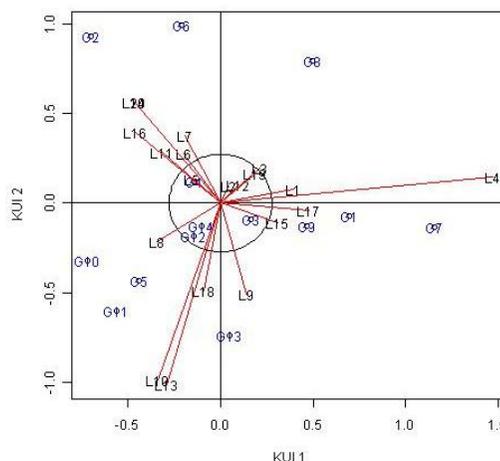
Kode	Lingkungan	Kode	Lingkungan
L1	Banteang1	L11	Tabanan2
L2	Narmada1	L12	Ngawi2
L3	Bantul1	L13	Narmada2
L4	Purworejo1	L14	Pesawaran2
L5	Tabanan1	L15	Probolinggo2
L6	Ngawi1	L16	Purworejo2
L7	Pusakanagara1	L17	Pusakanagara2
L8	Pasar miring1	L18	Rangkasbitung2
L9	Asahan1	L19	Takalar2
L10	Bantul2	L20	Taman Bogo2

Kestabilan Genotipe pada Model AMMI

Interaksi genotipe \times lingkungan merupakan keragaman yang disebabkan oleh efek gabungan dari genotipe dan lingkungan (Dickerson 1962 dalam Kang 2002). Interaksi genotipe \times lingkungan berakibat adanya perbedaan rata-rata hasil panen tanaman padi dari genotipe yang ditanam pada lingkungan yang berbeda. Interaksi genotipe \times lingkungan berhubungan dengan daya adaptasi dan kestabilan dari suatu genotipe. Genotipe yang berinteraksi dengan lingkungan kurang memiliki daya adaptasi terhadap lingkungan dan dikatakan cenderung tidak stabil, sedangkan genotipe yang tidak berinteraksi dengan lingkungan memiliki daya adaptasi yang tinggi dan dikatakan cenderung stabil. Interaksi genotipe \times lingkungan dapat diuraikan menjadi komponen utama interaksi dan kestabilan genotipe dapat dilihat dari nilai keragaman komponen utamanya.



(a)



(b)

Gambar 2 Biplot $AMMI_1$ (plot KUI_1 terhadap rata-rata hasil padi) (a) dan Biplot $AMMI_2$ (Plot KUI_1 terhadap KUI_2) (b).

Nilai singular dari matriks intraksi menyatakan bahwa banyaknya komponen yang dapat dipertimbangkan untuk model AMMI adalah komponen ke-1 sampai komponen ke-13, namun hanya 7 komponen utama interaksi yang berpengaruh nyata terhadap model. Kontribusi keragaman pengaruh interaksi yang mampu diterangkan oleh sembilan komponen adalah 23.11%, 21.34%, 17.50%, 9.92%, 7.48%, 6.04% dan 4.32%.

Biplot antara KUI_1 dengan rata-rata hasil sebagai biplot $AMMI_1$ merupakan tampilan grafis yang meringkas informasi pengaruh utama genotipe dan lingkungan yaitu pada sumbu rataan dan pengaruh interaksi genotipe dengan lingkungan pada sumbu KUI_1 . Genotipe yang letaknya satu titik pada

sumbu datar berarti mempunyai pengaruh utama yang sama dan jika terletak pada satu titik pada sumbu utama tegak berarti mempunyai pengaruh interaksi yang sama (Sumertajaya 2005).

Biplot $AMMI_1$ ditunjukkan pada Gambar 2(a). Biplot $AMMI_1$ memperlihatkan bahwa G2 memiliki rata-rata hasil padi yang paling rendah dan G12 memiliki rata-rata hasil yang paling tinggi. G1, G6 dan G8 mempunyai rata-rata hasil panen yang sama, namun mempunyai interaksi berbeda pada L9. G4, G12 dan G14 mempunyai rata-rata hasil padi dan keragaman KUI_1 yang sama.

Gambar 2(b) merupakan biplot $AMMI_2$ dengan daerah kestabilan ellipsnya. Hasil plot ini dapat menggambarkan keragaman interkasi sebesar 44.45%, masing-masing 23.11% untuk KUI_1 dan 21.34% untuk KUI_2 . Berdasarkan selang kepercayaan komponen utama interaksinya, genotipe yang dikategorikan stabil adalah G3, G4, G12 dan G14. Genotipe yang paling tidak stabil adalah G2 dan G7 dengan keragaman yang paling besar. G6, G8, G13, G11 dan G10 memiliki keragaman yang relatif besar, sehingga dapat dikategorikan bahawa genotipe tersebut tidak stabil. Lingkungan yang memiliki keragaman paling besar adalah L4, L10 dan L13.

Kestabilan Genotipe Model AMMI dengan Metode Resampling Bootstrap

Biplot antara KUI_1 dan KUI_2 yang digunakan untuk melihat kestabilan genotipe pada model AMMI hanya berupa suatu titik dari setiap genotipe yang menyatakan keragaman dari genotipe tersebut. Pengklasifikasian genotipe stabil yang berdasarkan satu titik dapat menimbulkan keragu-raguan bagi peneliti. Oleh karena itu dapat dilakukan pendekatan metode resampling bootstrap untuk meningkatkan keyakinan dalam menentukan kestabilan suatu genotipe. Metode resampling bootstrap digunakan untuk menentukan daerah kepercayaan skor KUI_1 dan KUI_2 pada biplot.

Untuk memperoleh galat baku dari skor KUI_1 dan KUI_2 dilakukan simulasi data hasil ulangan untuk memperoleh matriks interaksi. Resampling bootstrap dilakukan pada data asli, yaitu dengan cara mengambil 3 data baru dari 3 data ulangan dengan pemulihan untuk setiap genotipe di setiap lingkungan. Untuk setiap genotipe di setiap lingkungan terdapat $3^3 = 27$ kombinasi data yang mungkin terjadi. Dari 14 genotipe di 20 lingkungan dengan 3 kali ulangan diperoleh $27^{14 \times 20}$ kombinasi data yang mungkin terjadi.

Resampling bootstrap dilakukan dengan iterasi sebanyak 1000 kali. Di setiap iterasi diperoleh skor KUI_1^* dan KUI_2^* dari kombinasi data baru yang

Model AMMI Kini dan Yang Akan Datang

terbentuk. Setiap titik skor KUI_1^* dan KUI_2^* pada suatu genotipe dihitung jarak relatif r_g^* terhadap titik pusat (0,0), sehingga setiap genotipe memiliki galat baku $\widehat{se}_{r_g^*}$ yang berbeda sesuai dengan r_g^* . r_g^* dan $\widehat{se}_{r_g^*}$ masing-masing diperoleh dengan rumus berikut

$$r_{gi}^* = \sqrt{(KUI_{1i}^*)^2 + (KUI_{2i}^*)^2}, g = 1, 2, \dots, a; i = 1, 2, \dots, B^*$$

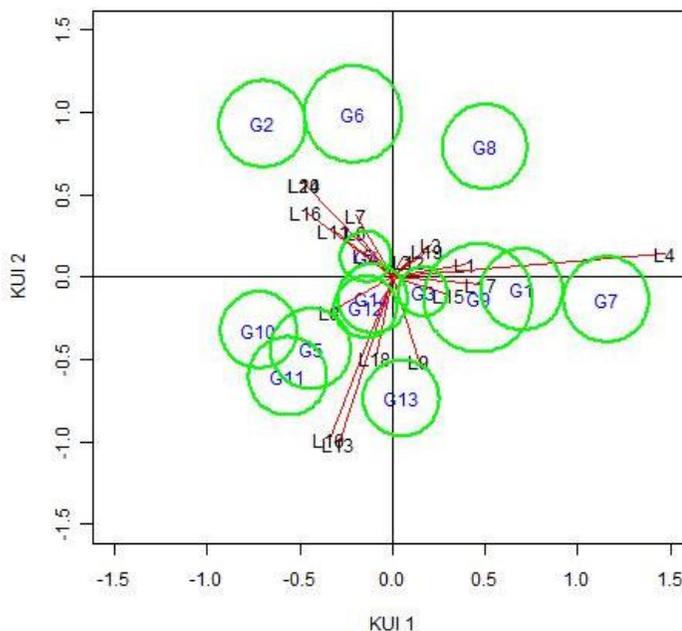
dan

$$\widehat{se}_{r_g^*} = \left\{ \frac{\sum_{i=1}^B [r_{gi}^* - \bar{r}_g^*]^2}{B-1} \right\}^{1/2} \quad (2)$$

dimana

$$\bar{r}_g^* = \frac{\sum_{i=1}^B r_{gi}^*}{B}$$

Galat baku dari setiap genotipe yang diperoleh pada pers.2 digunakan untuk membuat daerah kepercayaan skor KUI_1 dan KUI_2 . Daerah kepercayaan skor KUI_1 dan KUI_2 dibuat dalam bentuk lingkaran dengan galat baku $\widehat{se}_{r_g^*}$ sebagai jari-jari serta skor KUI_1 dan KUI_2 menjadi titik pusatnya. Daerah kepercayaan skor KUI_1 dan KUI_2 setiap genotipe digambar pada biplot $AMMI_2$ seperti yang terlihat pada gambar 3.



Gambar 3 Daerah kepercayaan skor KUI_1 dan KUI_2 menggunakan resampling bootstrap.

Gambar 3 merupakan hasil metode resampling bootstrap yang dapat dibedakan berdasarkan posisi dan luas lingkarannya. Berdasarkan posisinya, lingkaran suatu genotipe yang di dalamnya terdapat titik pusat (0,0) dikatakan sebagai genotipe stabil, sedangkan genotipe yang di dalam daerah lingkarannya terdapat titik skor KUI_1^* dan KUI_2^* suatu lingkungan tertentu dikatakan sebagai genotipe yang spesifik lingkungan. Berdasarkan luas daerah lingkarannya, suatu genotipe dikatakan semakin stabil apabila luas daerah lingkarannya semakin kecil.

Lingkaran dari suatu genotipe yang semakin mendekati dengan titik pusat (0,0) menunjukkan bahwa genotipe tersebut mempunyai galat baku skor KUI_1^* dan KUI_2^* yang relative tidak besar. Jarak r_g^* setiap titik skor KUI_1^* dan KUI_2^* untuk genotipe stabil mendekati nilai 0 dan dapat dikatakan bahwa genotipe tersebut dapat beradaptasi dengan baik di setiap lingkungan apapun. Genotipe spesifik lingkungan juga akan mempunyai jarak r_g^* dengan titik pusat (KUI_{1j}^*, KUI_{2j}^*) mendekati nilai 0 dan dikatakan genotipe tersebut dapat beradaptasi dengan baik di lingkungan tersebut saja.

Luas lingkaran yang besar diakibatkan oleh galat baku $\widehat{se}_{r_g^*}$ yang besar. Besarnya nilai galat baku $\widehat{se}_{r_g^*}$ sebagai jari-jari lingkaran daerah kepercayaan mengindikasikan bahwa simpangan baku skor KUI_1^* dan KUI_2^* hasil resampling bootstrap semakin beragam. Simpangan baku data asli yang besar akan menghasilkan galat baku $\widehat{se}_{r_g^*}$ data contoh bootstrap yang besar pula.

Pada gambar 3 genotipe yang di dalam lingkarannya terdapat titik pusat (0,0) adalah G4 dan G14 sehingga kedua genotype ini dikatakan genotipe yang stabil. Berdasarkan luas daerah kepercayaan, G4 mempunyai daerah yang lebih kecil daripada G14. Hal ini berarti ragam dari skor KUI_1^* dan KUI_2^* G4 lebih kecil dibandingkan ragam G14, sehingga G4 dapat dikatakan lebih stabil. Pada penelitian ini tidak terdapat genotipe yang spesifik lingkungan, namun ada beberapa genotipe yang berada di sekitar skor KUI_1 dan KUI_2 lingkungan tertentu. G5, G10 dan G11 relatif lebih dekat dengan L8. G13 relatif lebih cocok jika di tanam di lingkungan L9 dan L18, sedangkan G2 dan G6 lebih cocok ditanam di L20. Pada L4 genotipe yang lebih cocok ditanam adalah G1 dan G7. Genotipe yang tidak stabil berdasarkan luas daerah kepercayaan skor KUI_1^* dan KUI_2^* adalah G6, G7 dan G9, dimana genotipe ini berada di peringkat genotipe dengan luas daerah tiga terbesar utama.

4. Kesimpulan

Daerah kepercayaan rata-rata skor KUI_1^* dan KUI_2^* untuk metode resampling bootstrap dipengaruhi dan sangat sensitif terhadap posisi titik skor KUI_1 dan KUI_2 , sehingga jarak posisi skor KUI_1^* dan KUI_2^* akan lebih baik digunakan untuk menentukan kestabilan genotipe dibandingkan dengan menggunakan daerah kepercayaan skor KUI_1^* dan KUI_2^* . Menggunakan metode resampling bootstrap dapat diperoleh daerah kepercayaan skor KUI_1^* dan KUI_2^* setiap genotipe. Analisis AMMI menggunakan resampling bootstrap cocok untuk mendeteksi kestabilan pada percobaan multilokasi apabila ragam genotipe pada biplot kecil. Hasil kestabilan analisis AMMI menggunakan resampling bootstrap pada penelitian ini adalah G4 (OBS 1735/PSJ) dan G14 (CIMELATI).

Ucapan Terima Kasih

Terima kasih disampaikan kepada Direktorat Pendidikan Tinggi yang mendanai Hibah Penelitian Pascasarjana “Pengembangan Model Aditif untuk Seleksi Daya Adaptasi Tanaman” No.1/13.24.4/SPK/PD/2010.

Daftar Pustaka

- Chernick MR. (2008). *Bootstrap Methods: a Guide for Practitioners and Researches*, 2nd ed. New Jersey: John Wiley & Son, Inc.
- Cornelius PL. (1993). Statistical test and retention of terms in the additive main effects and multiplicative interaction model for cultivar trials. *Crop sci.* 33:1186-1193.
- Efron B, Tibshirani R. (1993). *An Introduction to the Bootstrap*. New York: Chapman & Hall.
- Hadi AF. (2006). Model AMMI terampat untuk data berdistribusi bukan normal [Tesis]. Bogor: Sekolah Pasca Sarjana, Institut Pertanian Bogor.
- Kang MS. (2002). Genotype-Environment Interaction: Progress and Prospects. Di dalam : Kang MS, Editor. *Quantitative Genetics, Genomic and Plant Breeding*. Florida: CRC Pr. Hlm .221-243.
- Lavoranti OJ, et.al. (2007). Phenotypic stability via ammi model with bootstrap re-sampling. *Boletim de Pesquisa Florestal.* 2:45-52. http://www.lce.esalq.usp.br/tadeu/Osmir_Tadeu_Wojtek.pdf, [5 februari 2010].

Mattjik AA, Sumertajaya IM. (2006). *Perancangan Percobaan dengan Aplikasi SAS dan Minitab Jilid 1, Edisi 2*. Bogor: IPB Press.

Sumertajaya IM. (2005). kajian pengaruh inter blok dan interaksi pada uji lokasi ganda dan respon ganda [disertasi]. Bogor: Sekolah Pasca Sarjana, Institut Pertanian Bogor.

Yuliyanti R. (2009). Identifikasi genotipe yang memberikan kontribusi terhadap interaksi genotipe x lingkungan pada model AMMI[Tesis]. Bogor: Sekolah Pasca Sarjana, Institut Pertanian Bogor.

Halaman ini sengaja dikosongkan

4

PENANGANAN DATA HILANG PADA MODEL AMMI

Dalam perancangan percobaan sering dijumpai data tidak lengkap yang disebabkan oleh banyak faktor. Data tidak lengkap yang dimaksudkan dalam kajian ini bisa pengamatan tidak lengkap atau bisa juga perlakuan yang tidak lengkap. Faktor-faktor yang berpotensi menyebabkan ketidaklengkapan data percobaan antara lain gangguan alam, atau bisa juga karena disengaja untuk efisiensi terutama percobaan-percobaan yang berskala besar.

Pada kasus data yang tidak lengkap ada dua cara untuk menganalisis data yaitu dengan menganalisis data yang teramati saja atau melakukan pendugaan terhadap data hilang lalu dilanjutkan dengan analisis datanya. Khusus dalam analisis AMMI kelengkapan data merupakan hal yang sangat penting, jika data tidak lengkap tidak diduga maka harus ada baris dan kolom yang dikorbankan untuk menyeimbangkan data. Pengorbanan data yang telah dikumpulkan bukan

merupakan strategi yang baik karena akan menghilangkan informasi dari beberapa perlakuan yang lain.

Seperti halnya percobaan-percobaan umum, ada beberapa permasalahan yang sering muncul dalam uji lokasi ganda. Permasalahan yang sering ditemui antara lain data yang dikumpulkan tidak lengkap (ada beberapa data pengamatan tidak teramati atau beberapa perlakuan tidak dicobakan karena berbagai kondisi). Karena berbagai kondisi alam (bencana alam, binatang, dll) mengakibatkan terjadinya kerusakan pada beberapa petak percobaan. Hal ini menyebabkan hilangnya informasi dari petak percobaan sehingga data yang diperoleh menjadi tidak lengkap.

Untuk mengatasi hal ini dalam analisis AMMI telah dikembangkan metode EM-AMMI (*Expectation Maximization AMMI*) untuk menduga data yang hilang.

4.1 EM-AMMI

Pendugaan data hilang dengan EM-AMMI dilakukan melalui proses iterasi. Pada tahap awal model AMMI dibangun berdasarkan data yang ada, kemudian data hilang diduga berdasarkan model AMMI yang diperoleh. Langkah selanjutnya adalah mengoptimalkan nilai dugaan dengan melakukan proses iterasi mulai tahap membangun model dan pendugaan data hilang berdasarkan model yang diperoleh. Proses ini dilakukan sampai data dugaan tidak berubah (konvergen).

Metode lain yang juga dapat digunakan dalam menduga data tidak lengkap yaitu metode *connected data*, dalam metode ini dipersyaratkan bahwa data yang teramati dapat dihubungkan melalui garis lurus yang tidak terputus atau sering juga disebut *g-connected*. Metode ini sebenarnya cukup baik digunakan sebagai nilai dugaan awal dari EM-AMMI.

4.2 Metode g-Connected

Menurut Weeks and Williams dalam Searle (1987), data pada tabel klasifikasi dua arah dikatakan terhubung secara geometri (*g-connected*) bila data pada sel-selnya dapat dihubungkan oleh garis lurus (garis horisontal maupun vertikal), yang mengikuti arah dari sel-sel terisi.

Struktur data yang terhubung secara geometri dimana setiap data yang ada dalam tabel dua arah tersebut dapat dihubungkan dengan suatu garis lurus yang tidak terputus. Sedangkan pada struktur data tak terhubung terdapat sel-sel

data yang ada terpisah menjadi dua kelompok data, dimana kedua kelompok data tersebut tidak dapat dihubungkan dengan suatu garis. Data yang tidak terhubung secara geometri biasanya memiliki beberapa kelompok data yang terhubung secara geometri, pendekatan analisis dapat dilakukan secara terpisah untuk masing-masing kelompok.

Untuk data *g*-connected berimplikasi bahwa seluruh dari kontras baris (μ_i) maupun kolom (τ_j) dapat diduga, hal ini berarti dalam tabel dua arah tanpa interaksi seluruh perbedaan $\mu_i - \mu_{i'}$, untuk $i \neq i'$, dan $\tau_j - \tau_{j'}$, untuk $j \neq j'$ dapat diduga (estimable). Akibatnya seluruh rata-rata sel dapat diduga, baik untuk sel-sel yang terisi maupun sel-sel yang kosong.

Untuk sel-sel terisi, pendugaan tidak menjadi masalah. Tetapi misal sel (*i,j*) kosong maka untuk data *g*-connected, ada beberapa sel (*i,j'*) dalam baris yang sama dan atau beberapa sel (*i',j*) dalam kolom yang sama yang terisi. Maka,

$$\mu_{ij} = \mu_i + \tau_j = \mu_i + \tau_j - \tau_{j'} + \tau_{j'} = \mu_{ij'} + (\tau_j - \tau_{j'})$$

karena kedua suku dapat diduga maka μ_{ij} dapat diduga.

Berikut adalah artikel tentang pendugaan data hilang pada model AMMI yang memanfaatkan *g*-connected dan EM-AMMI.

Halaman ini sengaja dikosongkan

PENDUGAAN DATA TIDAK LENGKAP DALAM AMMI

¹I Made Sumertajaya, ^{2,3}Gusti Ngurah Adhi Wibawa

¹Dosen Departemen Statistika Institut Pertanian Bogor

² Mahasiswa Pasca Sarjana Institut Pertanian Bogor

³ Staf Pengajar Jurusan Matematika F MIPA Universitas Haluoleo Kendari

ABSTRAK

Percobaan multilokasi mempunyai peranan penting dalam perkembangbiakan tanaman dan penelitian agronomi. Kajian mengenai interaksi antara genotipe dan lingkungan diperlukan dalam penyeleksian genotipe yang akan dilepas. Metode statistika yang biasa digunakan untuk mengolah data hasil percobaan multilokasi salah satunya adalah AMMI (Additive Main effect and Multiplicative Interaction). Metode ini menggabungkan analisis ragam aditif bagi pengaruh utama perlakuan dengan analisis komponen utama pada pengaruh interaksinya. Hambatan utama dalam menggunakan analisis AMMI adalah keseimbangan data. Sedangkan pada percobaan multilokasi kejadian data menjadi tidak seimbang peluangnya sangat besar. Agar setiap kombinasi genotipe dan lokasi memiliki jumlah ulangan yang sama, maka harus dilakukan pendugaan terhadap data yang tak lengkap. Pada kasus data tidak lengkap, diperlukan suatu metode pendugaan data untuk mempermudah analisis. Pada penelitian ini digunakan metode connected data dan algoritma EM-AMMI untuk menduga data yang tak lengkap dengan fungsi tujuannya adalah Mean Absolute Percentage Error (MAPE). Hasil simulasi menunjukkan bahwa semakin besar keragaman dari data maka semakin besar pula nilai MAPE semakin besar selaras juga dengan semakin tingginya persentas ketidaklengkapan data. Metode EM-AMMI memberikan MAPE yang lebih baik jika nilai awal dalam proses iterasi ditetapkan dengan menggunakan metode Connected. Sehingga metode Connected –EM AMMI merupakan metode yang layak untuk dijadikan dasar pendugaan data tidak lengkap dalam analisis AMMI.

Kata Kunci : AMMI, Connected AMMI, EM-AMMI

1. PENDAHULUAN

Pada setiap rancangan percobaan sering dijumpai data tidak lengkap yang disebabkan oleh banyak faktor. Data tidak lengkap yang dimaksudkan dalam kajian ini bisa pengamatan tidak lengkap atau bisa juga perlakuan yang tidak lengkap. Faktor-faktor yang berpotensi menyebabkan ketidaklengkapan data percobaan antara lain gangguan alam, atau bisa juga karena disengaja untuk efisiensi terutama percobaan-percobaan yang berskala besar.

Pada kasus data yang tidak lengkap ada dua cara untuk menganalisis data yaitu dengan menganalisis data yang teramati saja atau melakukan pendugaan terhadap data tidak lengkap lalu dilanjutkan dengan analisis datanya. Khusus dalam analisis AMMI kelengkapan data merupakan hal yang sangat penting, jika data tidak lengkap tidak diduga maka harus ada baris dan kolom yang dikorbankan untuk menyeimbangkan data. Pengorbanan data yang telah dikumpulkan bukan merupakan strategi yang baik karena akan menghilangkan informasi dari beberapa perlakuan yang lain.

Dalam analisis AMMI terdapat teknik pendugaan data tidak lengkap yaitu EM-AMMI, yang merupakan metode pendugaan data tidak lengkap secara iteratif yaitu menduga data tidak lengkap, mengepaskan model dan menduga kembali data tidak lengkap berdasarkan model yang diperoleh. Proses ini berhenti sampai kriteria konvergensi terpenuhi (Gauch, 1992). Untuk mendapatkan proses konvergensi lebih cepat maka diperlukan metode penentuan nilai awal yang lebih baik. Dalam kajian ini, nilai awal untuk data tidak lengkap akan dilakukan dengan metode connected data. Dalam metode connected data, data tidak lengkap diduga berdasarkan data yang teramati saja dengan menghubungkan perubahan kondisi disekitar data tidak lengkap. Dengan kombinasi metode connected data dengan algoritma EM-AMMI diharapkan pendugaan data tidak lengkap menjadi lebih teliti dan konvergensinya lebih cepat.

2. TINJAUAN PUSTAKA

Data Terhubungkan (Connected Data)

Menurut Weeks and Williams dalam Searle (1987), data pada tabel klasifikasi dua arah dikatakan terhubungkan secara geometri (*g-connected*) bila data pada sel-selnya dapat dihubungkan oleh garis lurus (garis horisontal maupun vertikal), yang mengikuti arah dari sel-sel terisi.

Gambar 1, menunjukkan struktur data yang terhubungkan secara geometri dimana setiap data yang ada dalam tabel dua arah tersebut dapat dihubungkan

dengan suatu garis lurus yang tidak terputus. Sedangkan pada Gambar2, terlihat bahwa data-data yang ada terpisah menjadi dua kelompok data, dimana kedua kelompok data tersebut tidak dapat dihubungkan dengan suatu garis sehingga struktur data pada Gambar2 termasuk salah satu contoh struktur data yang bukan g-connected.

	1	2	3	4	5
1	√		√		
2		√		√	
3			√		√
4	√			√	

Gambar 1. Struktur Data yang Bersifat G-connected

	1	2	3	4	5
1	√		√		
2		√		√	√
3			√		
4		√		√	

Gambar 2. Struktur Data yang Bersifat bukan G-connected

Data yang tidak terhubung secara geometri biasanya memiliki beberapa kelompok data yang terhubung secara geometri (kasus Gambar2, terdiri dari dua kelompok g-connected), pendekatan analisis dapat dilakukan secara terpisah untuk masing-masing kelompok.

Untuk data g-connected berimplikasi bahwa seluruh dari kontras baris (μ_i) maupun kolom (τ_j) dapat diduga, hal ini berarti dalam tabel dua arah tanpa interaksi seluruh perbedaan $\mu_i - \mu_{i'}$, untuk $i \neq i'$, dan $\tau_j - \tau_{j'}$, untuk $j \neq j'$ dapat diduga (estimable). Akibatnya seluruh rata-rata sel dapat diduga, baik untuk sel-sel yang terisi maupun sel-sel yang kosong.

Untuk sel-sel terisi, pendugaan tidak menjadi masalah. Tetapi misal sel (i,j) kosong maka untuk data g-connected, ada beberapa sel (i,j') dalam baris yang sama dan atau beberapa sel (i',j) dalam kolom yang sama yang terisi. Maka,

$$\mu_{ij} = \mu_i + \tau_j = \mu_i + \tau_j - \tau_{j'} + \tau_{j'} = \mu_{ij'} + (\tau_j - \tau_{j'})$$

karena kedua suku dapat diduga maka μ_{ij} dapat diduga.

EM-AMMI (Expectation - Maximization AMMI)

EM-AMMI ini merupakan perluasan metode AMMI dengan proses pendugaan (untuk data tidak lengkap) secara iteratif. Dalam penggunaan algoritma EM-AMMI ada beberapa hal yang membuat pendugaan ini tidak bisa dilakukan yaitu jika semua pengamatan hilang, pengamatan untuk satu genotipe atau satu lingkungan atau satu kelompok semua hilang, dan jika jumlah data tidak lengkap melebihi derajat bebas galat.

Ada dua tahap dasar dalam algoritma EM-AMMI yaitu :

1. Tahap E (Expectation) yaitu tahapan menduga data yang hilang.
2. Tahap M (Maximization) yaitu tahapan pembentukan model dari data yang ada.

Pada dasarnya metode ini bertujuan menduga model dengan data yang ada dengan mengabaikan data tidak lengkap. Lalu secara iteratif menduga data yang hilang kemudian dipaskan (dicari modelnya) hingga dicapai keadaan yang konvergen yaitu keadaan di mana perubahan nilai dugaan antar iterasi relatif kecil yaitu < 0.00001 (Gauch, 1992).

Jika pada analisis AMMI biasa skor KUI selalu tetap, misalnya skor KUI1 model AMMI1 sama dengan skor KUI1 pada model AMMI yang lebih tinggi, maka pada algoritma EM-AMMI skor KUI ini berubah pada tiap-tiap iterasi. Hal ini disebabkan nilai dugaan mempengaruhi nilai parameter aditif maupun multiplikatifnya.

Banyaknya iterasi untuk mencapai keadaan konvergen itu dipengaruhi oleh dua hal penting yaitu besar ukuran matriks dan persentase data tidak lengkap yang ada. Semakin besar ukuran matriks maka akan semakin lama pula iterasi yang diperlukan dan semakin besar persentase data tidak lengkap maka akan semakin besar pula jumlah iterasi yang dibutuhkan untuk mencapai kekonvergenan.

3. RANCANGAN SIMULASI DATA TIDAK LENGKAP

Untuk simulasi pendugaan data tidak lengkap dilakukan secara bertahap mulai dari penghilangan data lengkap lalu pemberian nilai awal dengan melihat pengaruh utamanya lalu menggunakan algoritma EM-AMMI untuk menghasilkan penduga bagi tiap-tiap data yang hilang. Data lengkap yang dihilangkan adalah data daya hasil jagung yang merupakan hasil percobaan multilokasi Jagung Hibrida yang dilakukan dari tanggal 23 Juli 2006 sampai 10 April 2007 yaitu pada musim hujan dan kemarau. Percobaan ini menggunakan 7 genotipe Jagung Hibrida Harapan dan 5 genotipe Jagung Hibrida Komersial. yang ditanam pada 18 lokasi tersebar di 6 Propinsi.

Secara ringkas algoritma simulasi pendugaan data tidak lengkap adalah sebagai berikut:

1. Penghilangan data secara acak tetapi tetap mempertahankan pola connected.
2. Pemberian nilai dugaan awal bagi data tidak lengkap menggunakan metode connected data. Pada tahap ini pengaruh genotif dan lokasi diasumsikan aditif, sehingga dugaan awal bagi data yang tidak lengkap dapat dilakukan sebagai berikut: (misalnya data tidak lengkap pada posisi genotif ke-2 dan lokasi ke-1, μ_{12})

Tabel 1. Struktur data pengamatan dengan perlakuan tidak lengkap

	L ₁	L ₂	...	L _b	Pengaruh genotif
G ₁	μ_{11}	μ_{12}		μ_{1b}	α_1
G ₂	Data tidak lengkap	μ_{22}		μ_{2b}	α_2
...		
G _a	μ_{a1}	μ_{a2}		μ_{ab}	α_a
Pengaruh lokasi	β_1	β_2		β_b	μ

Keterangan: μ_{ij} rata-rata genotif ke-i dan lokasi ke-j; α_i pengaruh genotif ke-i; β_j pengaruh lokasi ke-j; dan μ rata-rata umum.

Penduga bagi sel genotif ke-2 dan lokasi ke-1 adalah:

$$\hat{\mu}_{21} = \mu_{11} + (\alpha_2 - \alpha_1) \quad \text{atau} \quad \hat{\mu}_{21} = \mu_{22} + (\beta_1 - \beta_2)$$

$$\hat{\mu}_{21} = \bar{Y}_{11.} + (\bar{Y}_{2..} - \bar{Y}_{1..}) \quad \hat{\mu}_{21} = \bar{Y}_{22.} + (\bar{Y}_{.1.} - \bar{Y}_{.2.})$$

Penduga bagi masing-masing parameter dihitung berdasarkan data yang

diamati saja, sebagai berikut:

- Penduga bagi rata-rata genotif ke-i dan lokasi ke-j:

$$\hat{\mu}_{ij} = \frac{\sum_{k=1}^n Y_{ijk}}{n_{ij}} = \bar{Y}_{ij.},$$

- Penduga bagi pengaruh genotif ke-i: $\hat{\alpha}_i = \bar{Y}_{i...} - \bar{Y}_{...}$,
 - Penduga bagi pengaruh lokasi ke-j: $\hat{\beta}_j = \bar{Y}_{.j.} - \bar{Y}_{...}$
3. Pembangunan model AMMI, berdasarkan data yang sudah lengkap dengan tahapan sebagai berikut:
 - Penyusunan matriks dugaan pengaruh interaksi
 - Penguraian nilai singular matriks interaksi menjadi Komponen-komponen Utama Interaksi (KUI).
 - Penentuan banyaknya KUI dengan Metode *Posdictive Success*.
 - Pembentukan model AMMI.
 4. Memperbaiki nilai dugaan data tidak lengkap berdasarkan model AMMI terbaru.
 5. Pemeriksaan tingkat perubahan relatif nilai dugaan data tidak lengkap :
 - Jika tingkat perubahan $< 10^{-3}$ maka nilai dugaan baru tersebut merupakan nilai dugaan akhir.
 - Jika tingkat perubahan $> 10^{-3}$ maka ganti nilai dugaan dengan nilai dugaan baru dan ulangi langkah ke-3 sampai ke-5.
 6. Ketelitian nilai dugaan data tidak lengkap diukur dengan menggunakan *Mean Absolute Percentage Error* (MAPE), sebagai berikut:

$$MAPE = \sum_{i=1}^n \frac{\|Y_i - F_i\|}{Y_i} \times 100\%$$

Besaran nilai MAPE lebih besar dari 0%, semakin kecil nilai MAPE mengindikasikan nilai dugaan semakin teliti.

Banyaknya data tidak lengkap yang disimulasikan yaitu 5%, 10%, 20%, 30%, 40% dan 50%. Setiap gugus data yang dihilangkan akan dicoba 10 pola *connected*. Penghilangan data dilakukan dengan tetap mempertahankan syarat dari *connected data*.

4. HASIL SIMULASI DATA TIDAK LENGKAP

Pendugaan Data tidak lengkap Dengan Metode *Connected*

Pendugaan data tidak lengkap melalui metode *connected* secara teknis tidak terlalu sulit untuk dikerjakan. Data yang teramati dipersyaratkan terhubung

dalam rangkaian garis linier yang tak terputus dimaksudkan agar seluruh data yang hilang dapat dijelaskan dari kondisi data sekitarnya. Sebagai ilustrasi data pengamatan yang mengandung data tidak lengkap seperti terlihat pada **Tabel 2**.

Model linier aditif diasumsikan sebagai berikut: $y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$

Dimana μ rata-rata umum, α_i pengaruh genotif ke-i, β_j pengaruh lokasi ke-j, ε_{ij} pengaruh acak genotif ke-i dan lokasi ke-j.

Langkah pertama adalah menduga pengaruh genotif dan lokasi berdasarkan data yang teramati dengan metode kuadrat terkecil (*least square*) dengan restriksi jumlah pengaruh genotif dan lokasi sama dengan nol, sebagai berikut:

Pengaruh genotype ke-i:

$$\alpha_j = \bar{y}_{.j} - \bar{y}_{..}$$

$$\alpha_1 = 8.887 - 9.182 = -0.295$$

$$\alpha_2 = 9.021 - 9.182 = -0.161$$

dan seterusnya

Pengaruh lokasi ke-i:

$$\beta_i = \bar{y}_{i.} - \bar{y}_{..}$$

$$\beta_1 = 6.912 - 9.182 = -2.271$$

$$\beta_2 = 11.218 - 9.182 = 2.036$$

dan seterusnya

Hasil selengkapnya dapat dilihat pada **Tabel 2** kolom paling kanan untuk pengaruh genotif dan baris paling bawah untuk pengaruh lokasi.

Selanjutnya pendugaan data tidak lengkap dapat dilakukan sebagai berikut:

$$\mu_{ij} = \mu_{ij'} + (\beta_{j'} - \beta_j)$$

atau

$$\mu_{ij} = \mu_{i'j} + (\alpha_{i'} - \alpha_i)$$

Dengan demikian penduga data tidak lengkap di atas dapat diduga sebagai berikut:

Genotif F di lokasi L3.

$$\hat{y}_{3,1} = \hat{\mu}_{3,1} = \hat{\mu}_{2,1} + (\hat{\alpha}_3 - \hat{\alpha}_2) = 8.103 + (0.030 - (-1.147)) = 9.280$$

Hasil pendugaan selengkapnya dapat dilihat pada Tabel 2.

Tabel 2. Data Rata-Rata Daya Hasil yang Dihilangkan Sebanyak 5% yang Mengikuti Pola Terhubung (connected).

LOKASI	GENOTYPE											Rataan Lokasi	Pengaruh Lokasi	
	A	B	C	D	E	F	G	H	I	J	K			L
L2	7.590	6.333	6.193	6.420	7.140	7.440	6.977	6.370	7.529	6.677	7.073	7.137	6.912	-2.271
L3	10.194	10.165	11.313	13.100	13.131	.	10.527	10.897	11.646	10.250	.	10.439	11.218	2.036
L4	7.352	8.968	8.737	8.103	7.906	7.895	6.808	7.168	9.142	8.391	7.565	8.387	8.035	-1.147
L5	9.157	9.471	9.637	.	9.560	9.755	8.216	9.234	9.199	8.677	9.860	8.523	9.212	0.030
L6	8.011	9.584	7.668	8.274	8.415	8.301	7.332	.	8.698	9.563	7.758	8.854	8.405	-0.777
L7	7.591	9.105	7.549	9.173	8.728	7.471	8.931	7.943	8.584	6.388	8.851	8.626	8.248	-0.934
L8	7.783	7.790	8.834	10.426	9.355	9.038	9.534	8.800	9.187	8.303	9.108	8.783	8.915	-0.267
L9	9.614	10.623	10.706	.	11.460	11.401	12.041	12.163	10.207	10.415	10.414	9.910	10.814	1.631
L10	9.175	8.533	8.948	10.499	9.704	10.357	10.146	9.845	9.780	8.025	9.887	9.139	9.503	0.321
L11	10.245	9.438	10.284	.	11.087	10.940	9.660	11.048	10.665	.	9.913	10.932	10.421	1.239
L13	7.504	7.760	8.076	9.328	7.651	.	6.919	7.230	6.980	.	8.743	6.920	7.711	-1.471
L14	11.677	11.357	12.472	12.957	11.701	11.706	11.958	12.331	11.916	10.994	10.370	11.462	11.775	2.593
L15	10.765	9.441	9.354	10.084	10.935	10.748	10.299	10.197	10.611	10.758	10.473	10.158	10.318	1.136
L16	7.926	7.832	7.955	9.006	7.379	8.126	6.915	7.796	7.436	6.717	7.586	.	7.698	-1.485
L17	13.226	12.523	14.46	15.055	13.067	13.339	13.062	12.826	12.666	11.853	12.674	13.069	13.125	3.943
L18	4.374	5.415	4.671	5.679	4.341	5.286	4.435	5.148	4.632	4.684	4.435	5.275	4.859	-4.324
Rataan Genotype	8.887	9.021	9.183	9.813	9.510	9.413	8.986	9.281	9.311	8.692	8.995	9.176	9.182	
Pengaruh Genotype	-0.295	-0.161	0.001	0.631	0.327	0.231	-0.196	0.099	0.129	-0.490	-0.187	-0.006		

Besarnya nilai persentase simpangan mutlak (APE) bervariasi dari satu data ke data tidak lengkap yang lain. Hal ini mengindikasikan jika pengamatan yang hilang memiliki karakteristik yang jauh berbeda dengan nilai pengamatan disekitarnya maka nilai APE-nya akan cenderung besar sedangkan jika pengamatan yang hilang memiliki karakteristik yang mirip dengan nilai pengamatan disekitarnya maka nilai APE-nya akan cenderung kecil.

Rata-rata persentase simpangan mutlak (MAPE) untuk kelima buah data yang hilang adalah sebesar 8.21%. Dari nilai MAPE yang diperoleh dapat disimpulkan metode connected cukup baik menduga data yang hilang untuk pola yang pertama.

Dari 50 pola data yang dihilangkan untuk 5%, 10%, 20%, 30%, 40% dan 50% data yang hilang diperoleh nilai MAPE seperti pada **Error! Reference source not found.** Besarnya nilai MAPE terlihat cenderung membesar jika jumlah data yang dihilangkan semakin besar. Hal ini berarti jika semakin banyak data yang hilang maka akan menyebabkan tingkat ketelitian pendugaan akan menurun.

Tabel 3. Nilai Sebenarnya, Nilai Dugaan dengan Metode Connected dan Persentase Penyimpangan Relatif Antara Nilai Sebenarnya Dengan Dugaan

Lokasi	Genotype	Data Asli	Data Connected	Persentase Simpangan Mutlak
L3	F	11.984	11.746	1.984
L3	K	10.380	11.370	9.532
L5	D	10.708	9.280	13.342
L6	H	7.336	8.427	14.886
L9	D	11.533	12.324	6.861
L11	D	12.201	11.416	6.430
L11	J	10.716	8.943	16.549
L13	F	8.306	8.230	0.912
L13	J	6.594	6.233	5.478
L16	L	7.103	7.537	6.123
MAPE				8.209693

Tabel 4. Nilai MAPE untuk Pendugaan Data tidak lengkap Dengan Metode Connected

% Data Tidak Lengkap	CONNECTED
5	8.0427
10	8.1716
20	8.2429
30	8.2584
40	8.3186
50	8.3360

Pendugaan Data tidak lengkap dengan Metode EM-AMMI

Pendugaan data tidak lengkap dengan metode EM-AMMI membutuhkan nilai awal untuk data yang hilang. Dalam kajian ini nilai dugaan awal untuk EM-AMMI ditentukan sesuai dengan besaran dari dugaan metode connected. Hal ini dilakukan sekaligus untuk mengevaluasi apakah dugaan berdasarkan metode connected optimum berdasarkan metode EM-AMMI.

Model AMMI Kini dan Yang Akan Datang

Sebagai ilustrasi pendugaan data tidak lengkap seperti pada **Tabel 5**, selanjutnya akan diduga dengan metode EM-AMMI.

(i). Berikan nilai dugaan awal

$$\hat{y}_{26} = 11.747, \hat{y}_{211} = 11.370, \hat{y}_{44} = 9,280, \hat{y}_{58} = 8,427, \hat{y}_{84} = 12,324, \\ \hat{y}_{104} = 11,416, \hat{y}_{1010} = 8,943, \hat{y}_{116} = 8,230, \hat{y}_{1110} = 6,233, \hat{y}_{1412} = 7,537$$

(ii). Bangun model AMMI

Model linier AMMI yang digunakan adalah model AMMI tanpa ulangan sebagai berikut:

$$y_{ij} = \mu + \alpha_i + \beta_j + \sum_{k=1}^5 \sqrt{\lambda_k} \phi_{ki} \theta_{kj}$$

Dimana penduga masing-masing parameternya adalah sebagai berikut: Proses pendugaan ini konvergen pada iterasi ke-25, hasil pendugaan selengkapnya dapat dilihat pada **Tabel 5**.

Tabel 5. Nilai Sebenarnya, Nilai Dugaan Dengan Metode EM-AMMI dan Persentase Penyimpangan Relatif Antara nilai Sebenarnya Dengan Dugaan

Lokasi	Genotype	Data Asli	Data Connected-EM AMMI	Persentase Simpangan Mutlak
L3	F	11.984	11.649	2.799
L3	K	10.380	11.352	9.366
L5	D	10.708	9.908	7.471
L6	H	7.336	8.044	9.652
L9	D	11.533	12.209	5.864
L11	D	12.201	11.483	5.889
L11	D	12.201	11.483	5.889
L11	J	10.716	9.460	11.725
L13	F	8.306	7.893	4.978
L13	J	6.594	6.294	4.545
L16	L	7.103	7.516	5.814
MAPE				6.810

Rata-rata persentase simpangan mutlak (MAPE) untuk kelima buah data yang hilang adalah sebesar 6.810%. Dari nilai MAPE yang diperoleh dapat disimpulkan

untuk pola pertama, metode EM-AMMI cukup baik menduga data yang hilang untuk pola yang pertama bahkan hasilnya terlihat lebih teliti dibandingkan dengan metode connected.

Dari 50 pola data yang dihilangkan untuk 5%, 10%, 20%, 30%, 40% dan 50% data yang hilang diperoleh nilai MAPE seperti pada Tabel 6. Besarnya nilai MAPE dengan metode EM-AMMI juga terlihat cenderung membesar jika jumlah data yang dihilangkan semakin besar. Namun demikian sampai dengan 50% data tidak lengkap nilai MAPE masih dibawah 10% dan diperkirakan setelah data tidak lengkap melebihi 50% nilai MAPE akan lebih besar dari 10%

Tabel 6. Nilai MAPE untuk Pendugaan Data tidak lengkap Dengan Metode Connected-EM AMMI

% Data Tidak Lengkap	CONNECTED EM- AMMI
5	7.5159
10	7.5395
20	7.4140
30	7.4323
40	7.6707
50	7.7630

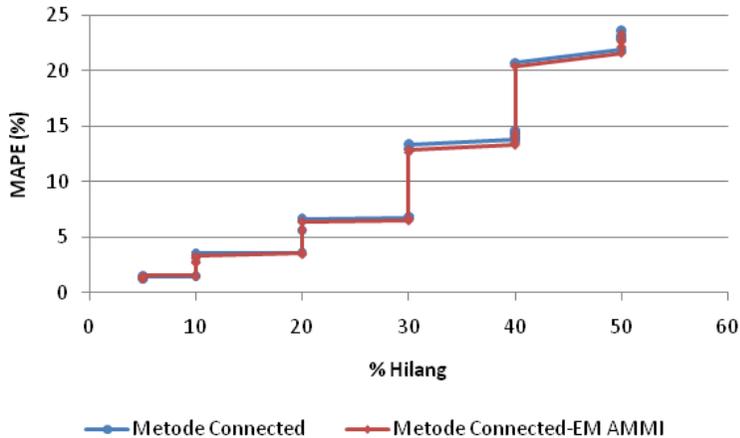
Selain menggunakan data asli, juga disimulasikan pendugaan data tidak lengkap dengan menggunakan data bangkitan. Analisa lebih lengkap mengenai data tidak lengkap disajikan di bawah ini :

Untuk simulasi data tidak lengkap dilakukan dengan menggunakan data bangkitan dan data asli yang beberapa persen datanya disimulasikan hilang. Di bawah ini merupakan hasil perbandingan pendugana data tidak lengkap menggunakan data hasil simulasi untuk $\varepsilon \sim (0,1)$, $\varepsilon \sim (0,5)$, $\varepsilon \sim (0,10)$, $\varepsilon \sim (0,20)$, dan $\varepsilon \sim (0,30)$.

Tabel 7. Perbandingan Pendugaan Data tidak lengkap dari Data Bangkitan

% Data Tidak Lengkap	Sigma ² (σ^2)	MAPE	
		CONNECTED	CONNECTED EM- AMMI
5%	1	1.268	1.311
	5	1.3292	1.3636
	10	1.4725	1.4614
	20	1.5009	1.4714
	30	1.5308	1.4981
10%	1	1.555	1.528
	5	3.2992	2.7587
	10	3.4378	2.8632
	20	3.5115	3.1636
	30	3.5848	3.3666
20%	1	3.6209	3.5136
	5	3.6556	3.6188
	10	5.6676	5.6968
	20	6.6196	6.4577
	30	6.6785	6.4253
30%	1	6.7905	6.5387
	5	6.8436	6.6974
	10	6.8823	6.6993
	20	12.9824	12.6782
	30	13.3856	12.9261
40%	1	13.7749	13.3326
	5	14.0591	13.4968
	10	14.332	13.9338
	20	14.6253	14.3871
	30	20.7163	20.4327
50%	1	21.8649	21.6013
	5	22.8944	22.1557
	10	23.0911	22.612
	20	23.1065	22.7325
	30	23.6626	23.2874

Model AMMI Kini dan Yang Akan Datang



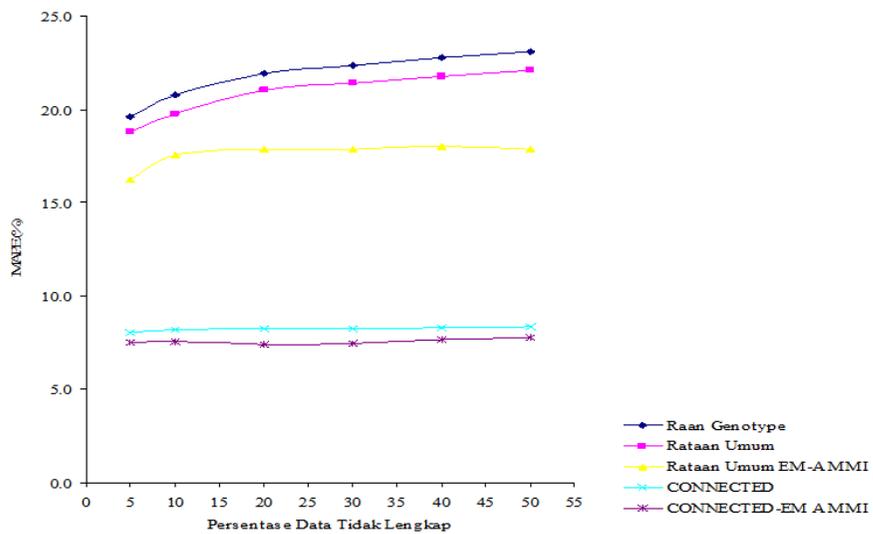
Gambar 3 Perbandingan Metode Pendugaan Data tidak lengkap Untuk Data Bangkitan

Gambar 3. memperlihatkan bahwa semakin besar keragaman dari data maka ada kecenderungan kelima metode pendugaan memberikan nilai MAPE yang semakin besar. Gambar di atas juga memberikan informasi bahwa dari kedua metode pendugaan data tidak lengkap. Metode Connected EM-AMMI memberikan pendugaan yang relatif paling akurat jika dibandingkan hanya dengan menggunakan metode Connected.

Tabel 8. Perbandingan Pendugaan Data tidak lengkap dari Data Daya Hasil Jagung

% Data Tidak Lengkap	MAPE				
	RATA-RATA GENOTYPE	RATA-RATA UMUM	EM-AMMI	CONNECTED	CONNECTED EM- AMMI
5%	19.6224	18.8177	16.2421	8.0427	7.5159
10%	20.7610	19.7802	17.5465	8.1716	7.5395
20%	21.9265	21.0552	17.8646	8.2429	7.4140
30%	22.3831	21.3948	17.8404	8.2584	7.4323
40%	22.7779	21.7634	18.0011	8.3186	7.6707
50%	23.1133	22.0911	17.8909	8.3360	7.7630

Dari Gambar 4. di atas terlihat bahwa metode Connected dan Connected-EM AMMI memberikan nilai MAPE paling rendah dan metode Connected-EM AMMI lebih baik dibandingkan metode Connected saja.



Gambar 4. Perbandingan Metode Pendugaan Data Tidak Lengkap Menggunakan Simulasi Data Daya Hasil Jagung

5. KESIMPULAN

Dari hasil yang diperoleh pada sub bab sebelumnya dapat ditarik kesimpulan sebagai berikut:

1. Pendugaan data tidak lengkap untuk ragam populasi yang semakin besar memberikan nilai MAPE yang semakin besar pula. Ini mengindikasikan bahwa semakin besar keragaman populasi metode pendugaan setiap metode pendugaan memberikan hasil yang kurang efisien.
2. Metode pendugaan data tidak lengkap yang paling efisien baik data hasil bangkitan ataupun data asli adalah metode Connected EM-AMMI. Metode Connected dapat memperbaiki efisiensi metode EM-AMMI dengan penetapan nilai awal yang lebih baik.
3. Semakin banyak jumlah data yang tidak lengkap mengakibatkan tingkat akurasi pendugaan semakin menurun. Banyaknya data yang tidak lengkap, yang masih dapat diduga dengan akurasi pendugaan cukup tinggi baik dengan menggunakan metode connected data maupun Connected EM-AMMI adalah maksimal 50%.

6. DAFTAR PUSTAKA

- Crossa, J. 1990. Statistical Analysis of Multilocation Trials. *Advances In Agronomy*. 44: 55-85.
- Johnson, R. A. and D.W. Winchern. 2002. *Applied Multivariate Statistical Analysis*. 5th ed. London: Prentice Hall International, Inc.

- Mattjik, A.A. 2000. Pendugaan Data tidak lengkap dengan Algoritma EM-AMMI pada Percobaan Lokasi Ganda. *Forum Statistika dan Komputasi, Vol. 5 No. 1*.
- Mattjik, A.A., dan I.M. Sumertajaya. 2002. *Perancangan Percobaan dengan Aplikasi SAS dan Minitab*. Bogor: IPB Press.
- Makridakis, S., S.C. Wheelwright, V.E. McGee. 1983. *Forecasting: Methods and Applications*. 2nd ed. Canada: John Wiley & Sons, Inc.
- Montgomery, D.C. 1991. *Design and Analysis of Experiments*. 3rd ed. New York: John Wiley & Sons, Inc.
- Searle, S.R. 1987. *Linear Models for Unbalance Data*. New York: John Wiley & Sons, Inc.

7. LAMPIRAN

Rataan Data Asli Daya Hasil Jagung

LOKASI	A	B	C	D	E	F	G	H	I	J	K	L
L2	7.59	6.33	6.10	6.49	7.24	7.44	6.92	6.37	7.58	6.68	7.06	7.14
L3	10.19	10.16	11.81	13.10	13.13	11.98	10.53	10.90	11.65	10.25	10.38	10.46
L4	7.35	8.97	8.74	8.10	7.91	7.89	6.81	7.17	9.14	8.39	7.57	8.39
L5	9.17	9.47	9.64	10.71	9.56	9.78	8.22	9.23	9.20	8.68	9.86	8.53
L6	8.01	9.58	7.67	8.27	8.41	8.30	7.33	7.34	8.70	9.56	7.76	8.85
L7	7.59	9.10	7.55	9.17	8.73	7.47	8.93	7.94	8.58	6.39	8.89	8.63
L8	7.78	7.79	8.85	10.43	9.36	9.04	9.55	8.80	9.19	8.30	9.11	8.79
L9	9.61	10.62	10.71	11.53	11.46	11.40	12.04	12.16	10.21	10.41	10.41	9.91
L10	9.18	8.53	8.95	10.50	9.70	10.36	10.15	9.84	9.78	8.02	9.89	9.14
L11	10.25	9.44	10.28	12.20	11.09	10.94	9.66	11.05	10.67	10.72	9.91	10.93
L13	7.50	7.76	8.08	9.33	7.65	8.31	6.92	7.23	6.98	6.59	8.74	6.92
L14	11.68	11.36	12.47	12.96	11.70	11.71	11.96	12.53	11.92	10.99	10.57	11.46
L15	10.76	9.44	9.35	10.08	10.94	10.75	10.30	10.20	10.61	10.76	10.47	10.16
L16	7.93	7.83	7.96	9.01	7.38	8.13	6.91	7.80	7.44	6.72	7.59	7.10
L17	13.23	12.52	14.15	15.05	13.07	13.34	13.06	12.83	12.67	11.85	12.67	13.07
L18	4.37	5.41	4.63	5.08	4.84	5.24	4.50	5.17	4.68	4.68	4.42	5.28

Program Metode Connected EM-AMMI

```
option nodate ps=120 ls=120;
Data EMAMMI;
Do i=13 To 180 By 1;
```

```
Output;
End;
PROC SURVEYSELECT DATA=EMAMMI
```

Model AMMI Kini dan Yang Akan Datang

```
METHOD=SRS n=10
      rep=10
      OUT=SAMPLEREP;

run;

data tidak lengkap;
set samplerep;
nk=12;
bagi=(i/nk);
sisa=mod(i,nk);
baris=int(bagi)+1;
if sisa=0 then do; kolom=nk;
baris=baris-1; end; else kolom=sisa;
drop sisa bagi;
run;

PROC IML;
load _all_;
use hilang;
read all;
b=(baris);
k=(kolom);
bk=(b||k);
n=10;
ul=10;
no=(1:n)~;
ulangan=(1:ul)~;
asli=repeat(0,n,1);
con=repeat(0,n,1);
mapec=repeat(0,ul,1);
mapem=repeat(0,ul,1);
em=repeat(0,n,1);
MEG={Matrik Data};

/*connected AMMI*/

sumasli=repeat(0,n,1);
sumcon=repeat(0,n,1);
sumem=repeat(0,n,1);
do i=1 to ul;
  program lengkap hubungi Penulis

/*EM-AMMI*/
  meg2=meg1;
  em0=con;

DO UNTIL (Y<0.0001);
  RG=MEG2[:,]; RE=MEG2[:,];
R=MEG2[:,];
  MRG=REPEAT(RG,16,1);
MRE=REPEAT(RE,1,12);
MR=REPEAT(R,16,12);
  MPEG=MEG2-MRE-MRG+MR;
CALL SVD(U,Q,V,MPEG);
F=Q##2; sumF=F[+,];
G=(F/SUMF)*100;
  s=0;
  do p=1 to 12;
    s=s+g[p,];
  A=p;
  if s>80 then p=13;

  program lengkap hubungi Penulis

PRINT no rataASLI rataCON rataEM;
print ulangan mapec mapem;
print 'Rata-rata MAPE';
print rmapec rmapem;
QUIT;
```

8. UCAPAN TERIMAKASIH

Kami Ucapkan Terimakaasih Kepada Direktorat Jenderal Pendidikan Tinggi (DIKTI). Tulisan ini bagian dari Hibah Penelitian Tim Pascasarjana yang didanai oleh Direktorat Jenderal Pendidikan Tinggi Nomor : 266/13.11/PL/2008 Tanggal : 02 April 2008.

5

MODEL AMMI UNTUK RESPON GANDA

Dalam mengkaji stabilitas genotipe dalam beberapa kasus tidak cukup dilihat hanya dari satu sisi. Selama ini kajian stabilitas hanya dilihat semata-mata berdasarkan produktifitas tanaman, hal ini terkesan kajian stabilitas masih terfokus pada masalah kuantitas belum menyentuh aspek kualitas (seperti bentuk, rasa, dan lain-lain). Oleh karena itu perlu pendekatan yang lebih komprehensif dalam melakukan kajian stabilitas yaitu dengan melibatkan beberapa respon yang memasukkan unsur kuantitas maupun kualitas.

Saat ini telah banyak berkembang strategi penggabungan peubah seperti metode range equalization, division by mean, first principal component (komponen

utama pertama) dan jarak Hotelling. Metode-metode ini banyak digunakan pada berbagai aspek seperti penyusunan indeks pembangunan manusia (human development index, HDI), indeks kemiskinan (poverty index, PI), penggabungan atribut ganda dalam analisis pengendalian mutu dan lain-lain. Pendekatan-pendekatan ini tentunya mempunyai kelemahan dan keunggulan pada kondisi-kondisi tertentu. Misalnya pendekatan *range equalization* maupun *division by mean* menerapkan sistem pembobotan yang sama pada seluruh peubah. Pendekatan ini tentunya akan sangat bagus bila tingkat kepentingan dari seluruh peubah dianggap sama, namun akan tidak baik bila peubah-peubah yang dilibatkan dalam kasus yang dihadapi memiliki tingkat kepentingan yang berbeda. Disamping itu, peubah-peubah yang dilibatkan dalam metode ini harus bersifat searah atau secara teoritis saling menunjang/menguatkan karena bobot dari semua peubah bertanda positif. Sedangkan pendekatan dengan komponen utama pertama cenderung lebih kuat mewakili peubah-peubah yang memiliki keragaman cukup besar, namun pendekatan ini akan sangat baik bila peubah-peubah yang dilibatkan memiliki korelasi yang cukup kuat.

Penggabungan respon merupakan salah satu strategi yang digunakan untuk menyederhanakan analisis untuk melihat daya adaptasi tanaman secara komprehensif. Strategi yang dapat diterapkan dalam permasalahan ini yaitu dengan membangun suatu peubah gabungan (indeks penampilan tanaman, IPT) berdasarkan beberapa respon tanaman. Beberapa pendekatan yang dapat digunakan yaitu: metode *range equalization* (IPT1), metode komponen utama pertama (IPT2), metode pembobotan berdasarkan komponen utama (IPT3), metode *Division by Mean* (IPT5) dan metode bobot optimum (IPT6).

5.1 Evaluasi Tingkat Validitas Metode Penggabungan Respon: Indeks Penampilan Tanaman (I Made Sumertajaya, G N A Wibawa, Seminar Nasional Matematika, UNY Yogyakarta, 22 Nopember 2008)

Evaluasi Tingkat Validitas Metode Penggabungan Respon: Indeks Penampilan Tanaman

¹Gusti N Adhi Wibawa, ²I Made Sumertajaya, ³Ahmad Ansori Mattjik

¹Mahasiswa S3 Pascasarjana Statistika IPB

^{2,3}Staf Pengajar Departemen Statistika Institut Pertanian Bogor

Abstrak

Selama ini analisis daya adaptasi tanaman lebih difokuskan hanya pada tingkat produksi atau daya hasil dan kurang memperhatikan aspek morfologi, daya tahan tanaman terhadap serangan hama dan penyakit maupun komponen hasil. Padahal diketahui bahwa setiap aspek-aspek tersebut merupakan komponen penting pada tanaman yang menjelaskan dimensi-dimensi yang berbeda. Sebagai ilustrasi, bobot 1000 butir mencerminkan dimensi bentuk dari butiran gabah, dimana jika bobot 1000 butir rendah dapat diartikan butiran gabah cenderung kurus-kurus sedangkan jika bobot 1000 butir tinggi dapat diartikan butiran gabah cenderung gemuk-gemuk. Dengan demikian analisis daya adaptasi tanaman yang hanya berfokus pada tingkat produksi semata tidak akan mendapatkan gambaran dari karakteristik tanaman secara komprehensif. Metode penggabungan respon merupakan salah satu metode yang dapat digunakan untuk menjawab permasalahan ini. Metode penggabungan respon dengan menggunakan sistem pembobotan berdasarkan komponen utama (IPT3) merupakan metode penggabungan respon yang terbaik, baik pada kelompok peubah berkorelasi rendah maupun tinggi.

PENDAHULUAN

Saat ini telah banyak berkembang strategi penggabungan peubah seperti metode *range equalization*, *division by mean*, *first principal component* (komponen utama pertama) dan jarak Hotelling. Metode-metode ini banyak digunakan pada berbagai aspek seperti penyusunan indeks pembangunan manusia (human development index, HDI), indeks kemiskinan (poverty index, PI), penggabungan atribut ganda dalam analisis pengendalian mutu dan lain-lain. Pendekatan-pendekatan ini tentunya mempunyai kelemahan dan keunggulan pada kondisi-kondisi tertentu. Misalnya pendekatan *range equalization* maupun *division by*

mean menerapkan sistem pembobotan yang sama pada seluruh peubah. Pendekatan ini tentunya akan sangat bagus bila tingkat kepentingan dari seluruh peubah dianggap sama, namun akan tidak baik bila peubah-peubah yang dilibatkan dalam kasus yang dihadapi memiliki tingkat kepentingan yang berbeda. Disamping itu, peubah-peubah yang dilibatkan dalam metode ini harus bersifat searah atau secara teoritis saling menunjang/menguatkan karena bobot dari semua peubah bertanda positif. Sedangkan pendekatan dengan komponen utama pertama cenderung lebih kuat mewakili peubah-peubah yang memiliki keragaman cukup besar, namun pendekatan ini akan sangat baik bila peubah-peubah yang dilibatkan memiliki korelasi yang cukup kuat.

Tujuan

Tujuan dari kajian ini adalah untuk memperbandingan tingkat validitas beberapa metode penggabungan respon tanaman dilihat dari kemampuan peubah gabungan mengikuti perilaku peubah asal. Langkah ini merupakan salah satu tahapan untuk mendapatkan metode penggabungan respon terbaik.

BAHAN DAN METODE

Bahan

Data yang digunakan dalam melakukan perhitungan dan evaluasi metode penggabungan respon (indeks penampilan tanaman, ipt) adalah dua gugus data simulasi. Gugus data pertama yaitu peubah-peubahnya berkorelasi rendah ($r < 0.5$) dan gugus data kedua yaitu peubah-peubahnya berkorelasi tinggi ($r \geq 0.5$). Setiap gugus data dibangkitkan sebanyak 100 kali.

Metode

Penggabungan respon merupakan salah satu strategi yang digunakan untuk menyederhanakan analisis untuk melihat daya adaptasi tanaman secara komprehensif. Respon gabungan yang diperoleh selanjutnya akan disebut sebagai indeks penampilan tanaman (IPT). Ada 6 metode pendekatan yang digunakan untuk menghitung IPT dari lima peubah respon yang diamati (tinggi tanaman, persen gabah isi, jumlah anakan produktif, bobot 1000 butir, dan daya hasil), yaitu metode *range equalization* (IPT1), metode komponen utama pertama (IPT2), metode pembobotan berdasarkan komponen utama (IPT3), metode jarak *Hotelling* (IPT4), metode *Division by Mean* (IPT5) dan metode bobot optimum (IPT6).

a. Pendekatan I (*Range Equalisation*)

Untuk memperoleh nilai respon gabungan digunakan informasi nilai minimum dan maksimum dari data respon peubah asal (Lawrence *et al*, 2003; Gani & Duncan, 2004; Kundu, 2004). Tahapan yang dilakukan untuk memperoleh nilai respon gabungan (IPT1) adalah:

1. Carilah nilai SDII (*sub dimension indicator index*) untuk masing-masing peubah asal, yaitu

$$SDII_i = \frac{Y_{ij} - Y_{i\min}}{Y_{i\max} - Y_{i\min}}$$

dengan $i = 1, 2, \dots, p$ dan $j = 1, 2, \dots, n$; p adalah banyaknya peubah asal dan n adalah banyaknya amatan

2. Carilah nilai IPT1 yaitu rata-rata dari seluruh SDII

$$IPT_1 = \sum_{i=1}^p \frac{SDII_i}{p}$$

b. Pendekatan II (Skor Komponen Utama Pertama)

Nilai respon gabungan (IPT2) yang diperoleh berdasarkan pendekatan ini merupakan nilai skor komponen utama pertama dari analisis komponen utama terhadap data respon peubah asal (Abeyasekera, 2005).

Tahapan analisis yang dilakukan pada pendekatan ini adalah sebagai berikut: Misal vektor peubah yang diamati adalah $\underline{Y}' = (Y_1, Y_2, \dots, Y_p)$

- i. Hitunglah matriks koragam (S) atau matriks korelasi (R)

$$S = \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1p} \\ S_{21} & S_{22} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ S_{p1} & \dots & \dots & S_{pp} \end{bmatrix} \text{ dimana } S_{ii} = \frac{\sum_{k=1}^n (y_{ik} - \bar{y}_i)^2}{n-1}$$

$$S_{ij} = \frac{\sum_{k=1}^n (y_{ik} - \bar{y}_i)(y_{jk} - \bar{y}_j)}{n-1}$$

Dan

$${}_p R_p = \begin{bmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ r_{p1} & \dots & \dots & 1 \end{bmatrix} \quad \text{dimana}$$

$$r_{ij} = \frac{\sum_{k=1}^n (y_{ik} - \bar{y}_i)(y_{jk} - \bar{y}_j)}{\sqrt{\sum_{k=1}^n (y_{ik} - \bar{y}_i)^2 \sum_{k=1}^n (y_{jk} - \bar{y}_j)^2}}$$

- ii. Carilah vektor ciri (*eigen vector*) dan akar ciri (*eigen value*) dari persamaan ciri berikut:

$$S\underline{a} = \lambda\underline{a} \text{ atau } R\underline{a} = \lambda\underline{a}$$

Dengan ketentuan sebagai berikut:

- ❖ Gunakan matriks koragam (S) jika peubah-peubah yang dianalisis memiliki satuan yang sama dan gunakan matriks korelasi (R) jika peubah-peubah yang dianalisis memiliki satuan yang berbeda
- ❖ Tata vektor ciri-vektor ciri $\underline{a}_1, \dots, \underline{a}_p$ yang berpadanan dengan akar ciri-akar ciri $\lambda_1 > \dots > \lambda_p$, dengan kendala $\underline{a}_i \underline{a}_i' = 1$ dan $\underline{a}_i \underline{a}_j' = 0$

- iii. Hitung IPT2 sebagai berikut:

IPT2 dapat dihitung menggunakan dua pendekatan sebagai berikut:

- Jika satuan peubah sama, $\text{IPT2} = \underline{a}_1' \underline{Y} = a_{11} Y_1 + a_{12} Y_2 + \dots + a_{1p} Y_p$
- Jika satuan peubah tidak sama, $\text{IPT2} = \underline{a}_1' \underline{Z} = a_{11} Z_1 + a_{12} Z_2 + \dots + a_{1p} Z_p$, dimana Z_i adalah peubah Y_i yang sudah dibakukan.

- c. Pendekatan III (Pembobotan berdasarkan Analisis Komponen Utama)

Besarnya bobot masing-masing respon akan dilakukan dengan pendekatan komponen utama. Tahapan penggabungan respon sama seperti tahapan yang dilakukan pada IPT2 tetapi respon gabungan tidak hanya ditentukan oleh komponen utama pertama melainkan dapat ditentukan oleh beberapa komponen utama. Banyaknya komponen utama yang dipilih ditentukan berdasarkan persentase keragaman kumulatif. Persentase keragaman kumulatif dapat dihitung sebagai berikut:

$$\text{Persentase keragaman komponen ke-} i = \frac{\lambda_i}{\sum_{j=1}^p \lambda_j} \times 100\%$$

$$\text{Persentase keragaman kumulatif } q \text{ komponen} = \frac{\sum_{j=1}^q \lambda_j}{\sum_{j=1}^p \lambda_j} \times 100\%$$

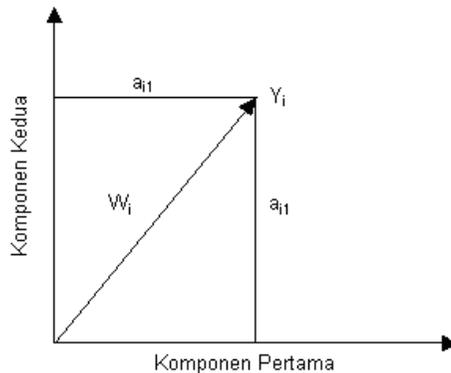
Batas minimal persentase keragaman kumulatif yang digunakan adalah 75%.
 Penentuan bobot dilakukan sebagai berikut: (kasus dua komponen)

$$Z_1 = a_{11} Y_1 + a_{12} Y_2 + \dots + a_{1p} Y_p$$

$$Z_2 = a_{21} Y_1 + a_{22} Y_2 + \dots + a_{2p} Y_p$$

Maka bobot untuk peubah ke- i diperoleh sebagai berikut: (lihat Gambar 4.1.)

$$W_i = \sqrt{\frac{a_{1i}^2}{\lambda_1} + \frac{a_{2i}^2}{\lambda_2}}$$



Gambar 1 Bobot peubah Y_i berdasarkan komponen utama pertama dan kedua

Sehingga respon gabungan (IPT3) adalah

$$\text{IPT3} = w_1 Y_1 + w_2 Y_2 + \dots + w_p Y_p$$

Bobot masing-masing peubah adalah mencerminkan besarnya keragaman peubah asal yang dijelaskan oleh komponen utama yang terpilih.

- d. Pendekatan IV (Jarak Hotelling)

Misal vektor peubah respon $\mathbf{x}'=(x_1, \dots, x_p)$, vektor rata-rata peubah respon $\bar{\mathbf{x}}'=(\bar{x}_1, \dots, \bar{x}_p)$ dan matriks kovarian S maka pendekatan Hotelling untuk objek ke-i dapat dirumuskan sebagai berikut:

$$d_i^2 = (\underline{x}_i - \bar{\underline{x}})' S^{-1} (\underline{x}_i - \bar{\underline{x}})$$

Nilai ini sering juga disebut jarak kuadrat Mahalanobis setiap titik pengamatan ke vector rata-rata. Jika antar peubah saling bebas maka pendekatan Hotelling dapat diubah menjadi:

$$d_i^2 = (\underline{x}_i - \bar{\underline{x}})' (\underline{x}_i - \bar{\underline{x}})$$

yang sering juga disebut jarak kuadrat Euclid.

Kemudian respon gabungan (IPT4) yang dimaksud adalah nilai d.

e. Pendekatan V (*Division by Mean*)

Untuk memperoleh nilai respon gabungan digunakan informasi nilai rata-rata dari data respon peubah asal (Kundu, 2004). Tahapan yang dilakukan untuk memperoleh nilai respon gabungan (IPT5) adalah:

1. Carilah nilai SDII untuk masing-masing peubah asal, yaitu

$$SDII_i = \frac{Y_{ij}}{\bar{Y}_i}$$

dengan $i = 1, 2, \dots, p$ dan $j = 1, 2, \dots, n$; p adalah banyaknya peubah asal dan n adalah banyaknya amatan

2. Carilah nilai IPT5 yaitu rata-rata dari seluruh SDII

$$IPT_5 = \sum_{i=1}^p \frac{SDII_i}{p}$$

Perhitungan dan evaluasi IPT dengan menggunakan data simulasi hanya dilakukan terhadap lima IPT selain IPT6. Tahapan proses perhitungan dan evaluasi IPT berdasarkan data hasil simulasi adalah:

- a. Bangkitkan data. Skenario pembangkitan data yang dibuat pada simulasi ini yaitu gugus data pertama peubah-peubahnya berkorelasi rendah yaitu tingkat korelasinya kurang dari 0.5, sedangkan gugus data kedua peubah-peubahnya berkorelasi tinggi yaitu tingkat korelasinya lebih besar atau sama dengan 0.5. Algoritmanya adalah sebagai berikut:
 - i. Bangkitkan 5 peubah (Y_1, Y_2, \dots, Y_5) dengan korelasi tertentu

- ii. Hitung respon gabungan dengan menggunakan kelima metode penggabungan peubah, yang selanjutnya disebut IPT1, IPT2, IPT3, IPT4 dan IPT5.
 - iii. Hitung korelasi setiap IPT dengan kelima peubah asal
 - iv. Simpan nilai mutlak korelasi minimum setiap IPT dengan kelima peubah asal.
- b. Evaluasi validitas masing-masing pendekatan (IPT) berdasarkan kedekatan IPT dengan peubah asal, yang didekati dengan nilai korelasi antara IPT dengan seluruh peubah asal. Tingkat validasi masing-masing IPT merupakan nilai korelasi minimum antara IPT dengan seluruh peubah asal. Kekonsistenan tingkat validasi masing-masing IPT akan dilihat dari hasil simulasi 100 kali pembangkitan data untuk setiap kelompok peubah (berdasarkan korelasi).

HASIL DAN PEMBAHASAN

Penggabungan respon merupakan salah satu strategi yang digunakan untuk menyederhanakan analisis untuk melihat daya adaptasi tanaman secara komprehensif. Respon gabungan yang diperoleh selanjutnya akan disebut sebagai indeks penampilan tanaman (IPT). Teknik penggabungan respon akan didekati dengan menggunakan lima metode yaitu metode *range equalisation* (IPT1), komponen utama pertama (IPT2), pembobotan berdasarkan komponen utama (IPT3), jarak *Hotelling* (IPT4), dan metode *division by mean* (IPT5).

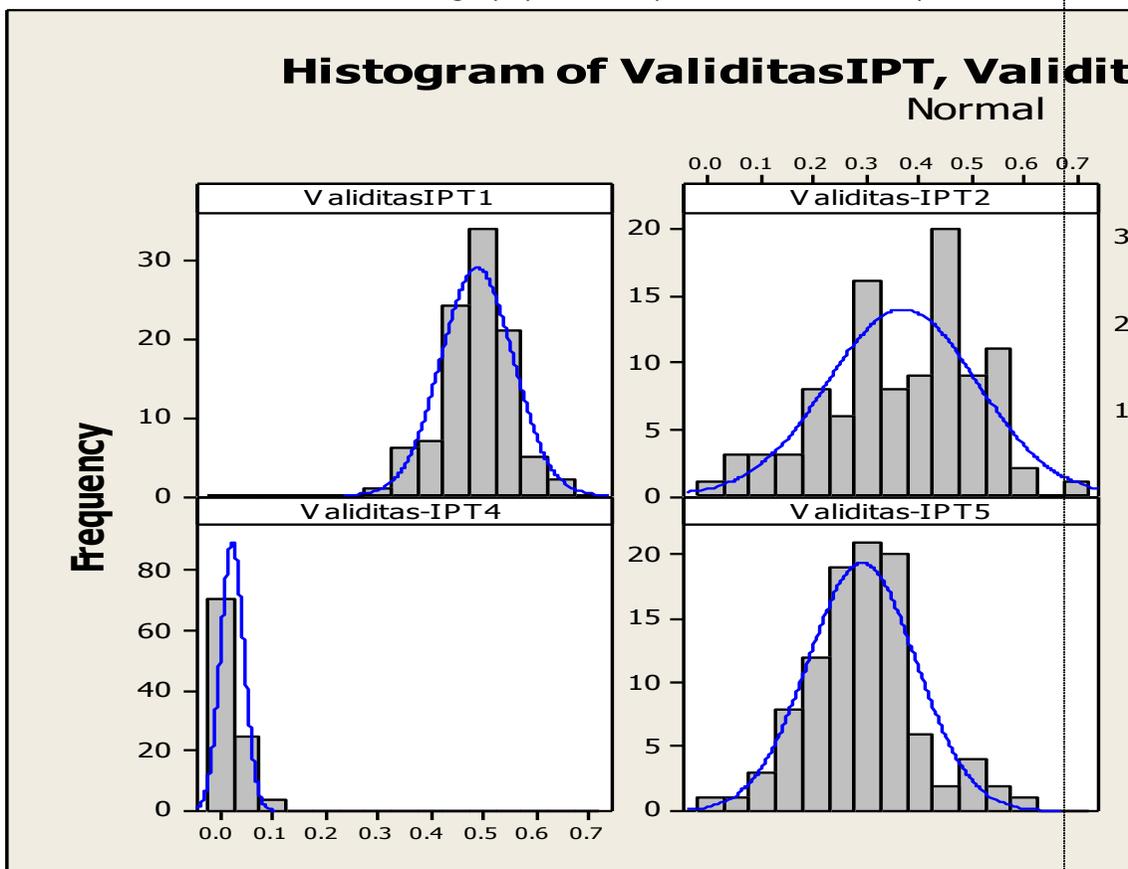
Validasi masing-masing pendekatan akan diukur berdasarkan kedekatan indeks dengan peubah asal, yang didekati dengan nilai korelasi antara indeks dengan seluruh peubah asal. Tingkat validasi masing-masing indeks merupakan nilai korelasi minimum antara indeks dengan seluruh peubah asal. Kekonsistenan tingkat validasi masing-masing indeks akan dilihat dari hasil simulasi 100 kali pembangkitan data untuk setiap kelompok peubah. Kelompok peubah pertama yang dibangkitkan peubah-peubah yang berkorelasi rendah yaitu dengan tingkat korelasi di bawah 0.5 dan kelompok peubah kedua yang dibangkitkan adalah peubah-peubah yang berkorelasi tinggi dengan tingkat korelasi di atas 0.5. Program simulasinya dapat dilihat pada **Lampiran 1**.

Kelompok peubah berkorelasi rendah

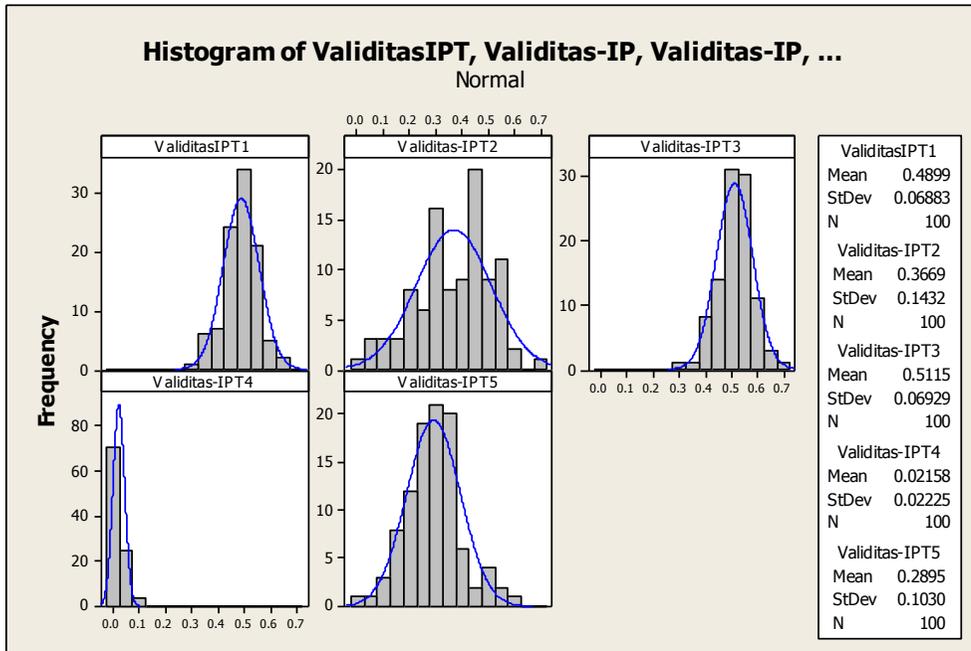
Model AMMI Kini dan Yang Akan Datang

Dari hasil 100 kali simulasi diperoleh metode penggabungan respon yang memiliki tingkat validitas tertinggi adalah metode pembobotan berdasarkan komponen utama (IPT3) yaitu dengan tingkat validitas rata-rata sebesar 0.5115, yang artinya korelasi minimum yang terjadi antara gabungan respon dengan seluruh peubah asal sebesar 0.5115. Sedangkan metode penggabungan respon yang juga memiliki tingkat validitas cukup besar yaitu metode range equalization (IPT1) yaitu sebesar 0.4899, yang disusul kemudian secara berturut-turut oleh metode komponen utama pertama (IPT2) dengan tingkat validitas sebesar 0.3669, metode *division by mean* (IPT5) dengan tingkat validitas sebesar 0.2895, dan metode jarak Hotelling memiliki tingkat validitas paling kecil yaitu sebesar 0.02158.

Hasil selengkapnya dapat dilihat pada

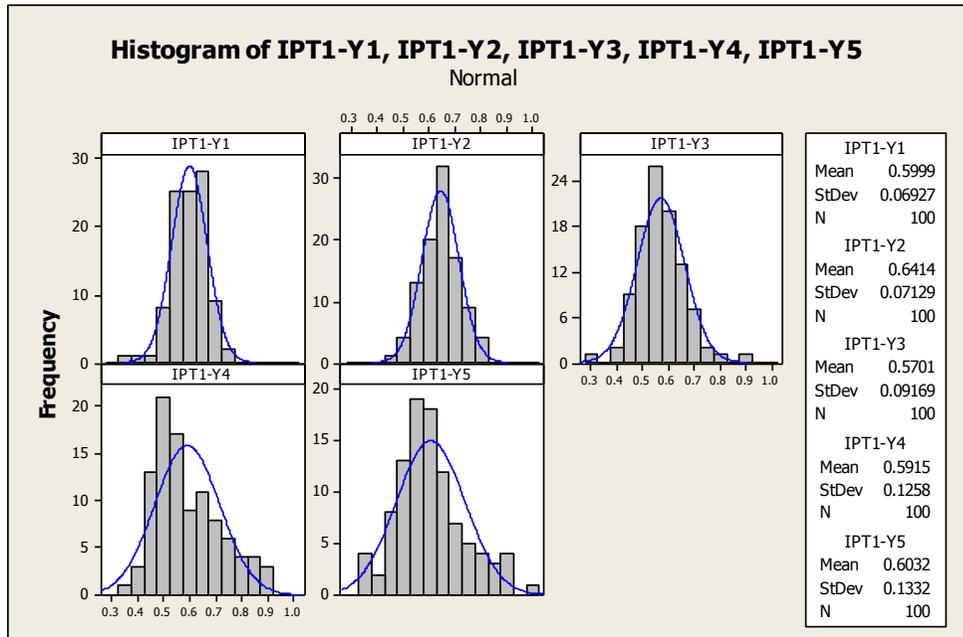


Gambar 2.



Gambar 2 Perbandingan tingkat validitas kelima metode penggabungan respon yang berkorelasi rendah dari 100 kali simulasi

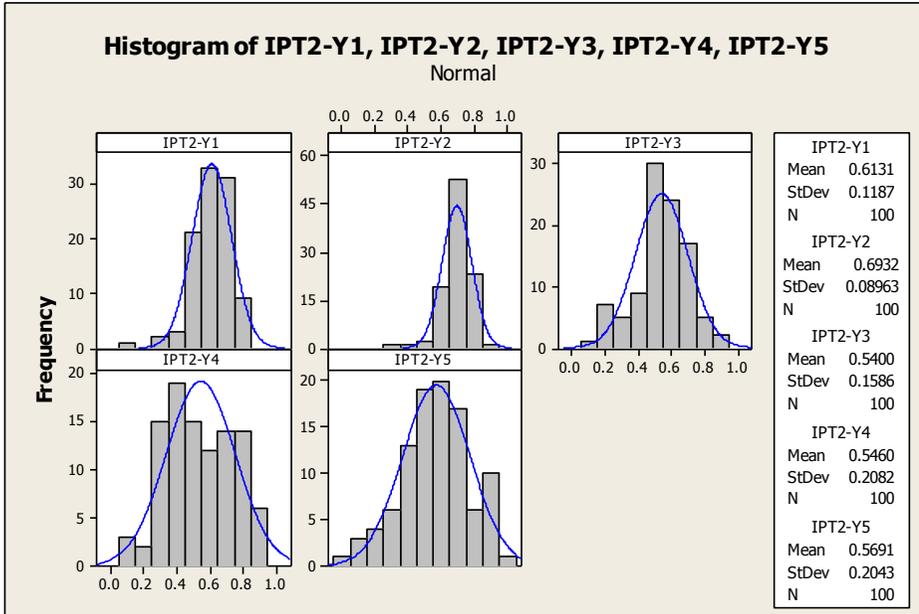
Hasil ini menunjukkan bahwa metode pembobotan berdasarkan komponen utama (IPT3) untuk sementara merupakan pilihan terbaik dalam menggabungkan respon-respon yang berkorelasi rendah. Namun demikian hasil ini harus diperkuat dengan melihat tingkat korelasi antara gabungan respon dengan setiap peubah asal. Gabungan respon yang baik diharapkan memiliki tingkat korelasi yang tinggi dengan seluruh peubah asal, jika hal ini terjadi berarti gabungan respon mampu menjadi wakil yang baik dalam menjelaskan perilaku seluruh peubah asal.



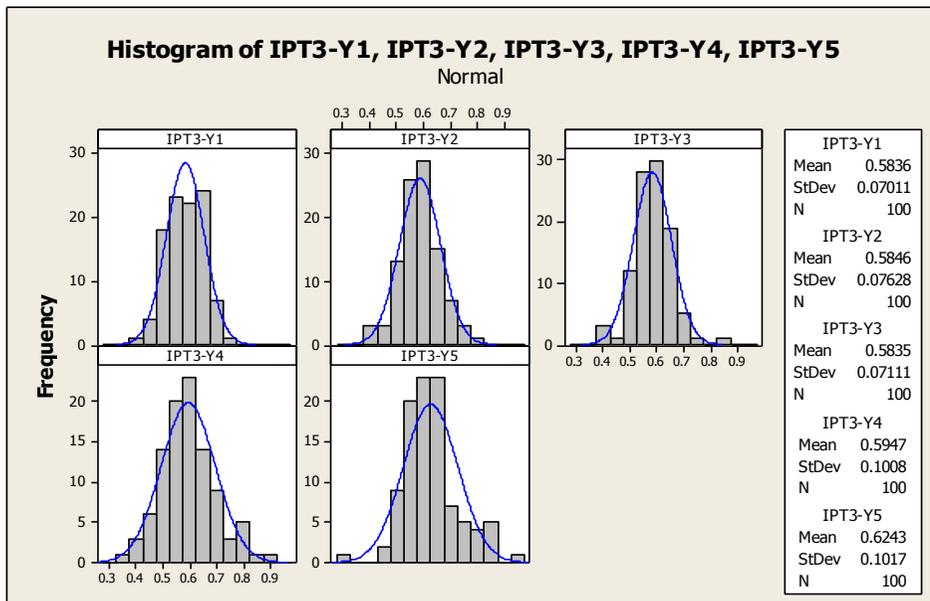
Gambar 3. Keragaan nilai mutlak korelasi antara gabungan respon dengan metode range equalization (IPT1) dengan seluruh peubah asal kelompok pertama dari 100 kali simulasi

Dari **Gambar 3** sampai **Gambar 7**, terlihat IPT1, IPT2, IPT3 dan IPT5 memiliki tingkat korelasi yang cukup besar dan merata dengan setiap peubah asal yaitu berkisar 0.5 sampai 0.7. Namun demikian jika dilihat dari kestabilan nilai korelasi yang dihasilkan, IPT3 adalah metode penggabungan yang terbaik. Hal ini terlihat dari nilai simpangan bakunya yang paling kecil yaitu berkisar antara 0.07 sampai 0.10. Kemudian secara berturut-turut diikuti oleh IPT1, IPT5 dan IPT2.

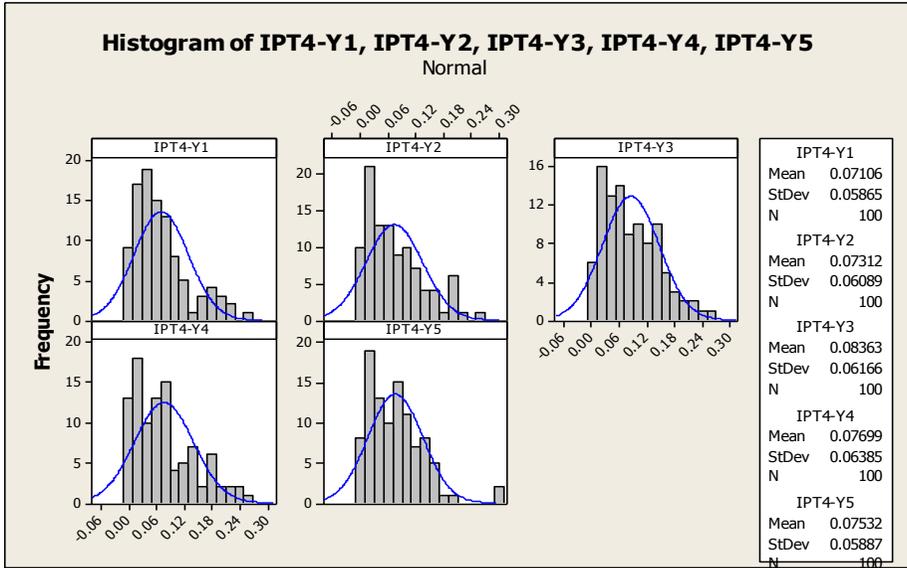
Perilaku yang berbeda ditunjukkan oleh IPT4 dimana tingkat korelasinya dengan peubah asal sangat kecil dibandingkan dengan keempat metode yang lain yaitu berkisar antara 0.07 sampai 0.08, tetapi besaran nilai korelasinya paling stabil dibandingkan dengan metode-metode yang lain yaitu dengan simpangan baku 0.5 sampai 0.6. Namun demikian IPT4 bukan merupakan perwakilan yang baik untuk peubah asal karena tingkat korelasinya sangat rendah sehingga IPT4 tidak mampu menjelaskan perilaku peubah-peubah asal dengan baik.



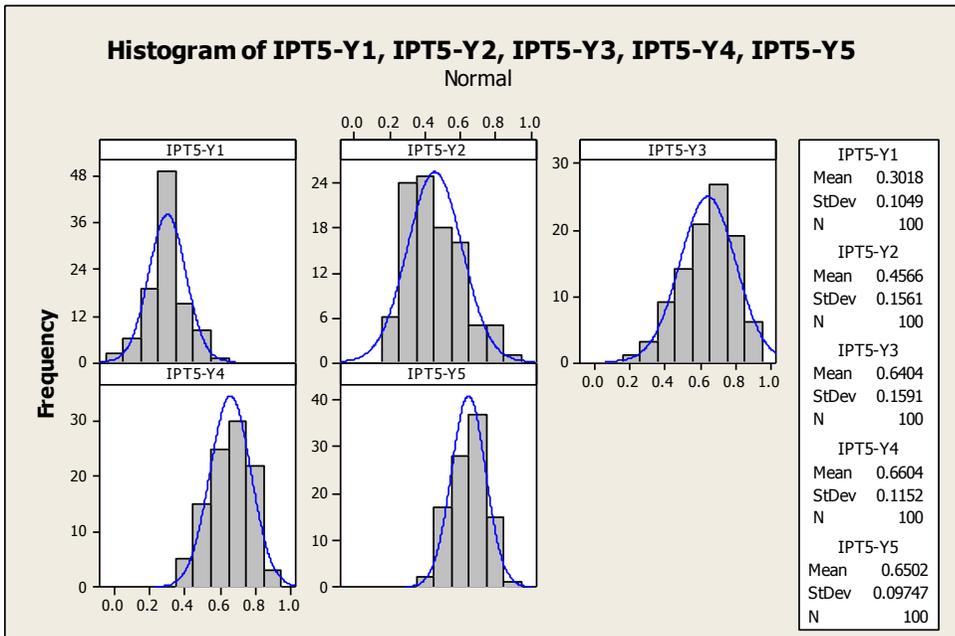
Gambar 4 Keragaan nilai mutlak korelasi antara gabungan respon dengan metode komponen utama pertama (IPT2) dengan seluruh peubah asal kelompok pertama dari 100 kali simulasi



Gambar 5. Keragaan nilai mutlak korelasi antara gabungan respon dengan metode pembobotan berdasarkan komponen utama (IPT3) dengan seluruh peubah asal kelompok pertama dari 100 kali simulasi



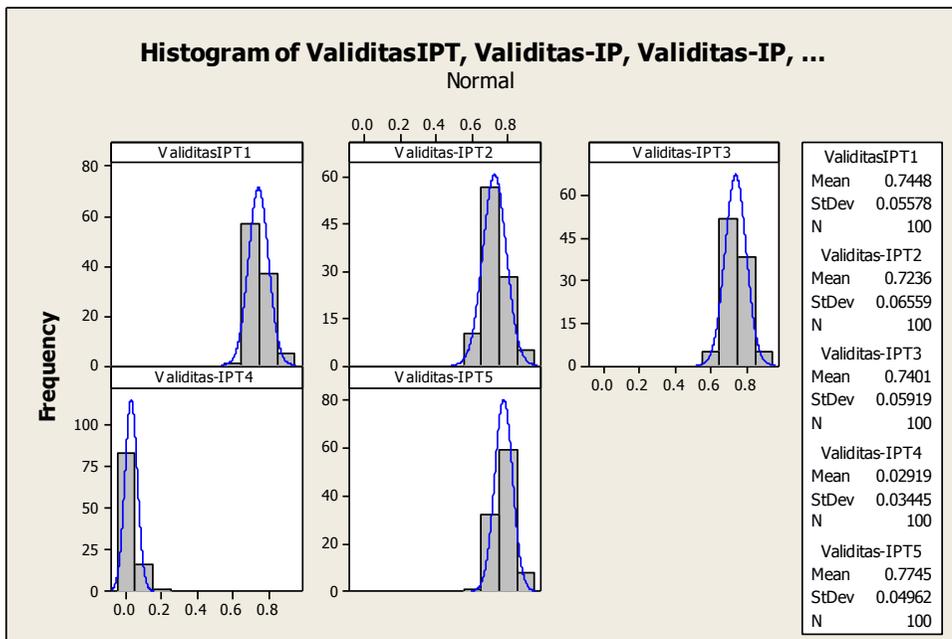
Gambar 6. Keragaan nilai mutlak korelasi antara gabungan respon dengan metode jarak Hotelling (IPT4) dengan seluruh peubah asal kelompok pertama dari 100 kali simulasi



Gambar 7. Keragaan nilai mutlak korelasi antara gabungan respon dengan metode division by mean (IPT5) dengan seluruh peubah asal kelompok pertama dari 100 kali simulasi

Kelompok peubah berkorelasi tinggi

Pada kelompok peubah berkorelasi tinggi dari hasil 100 kali simulasi diperoleh IPT1, IPT2, IPT3 dan IPT5 memiliki tingkat validitas yang hampir sama yaitu berkisar antara 0.72-0.77. Hasil ini agak berbeda dengan hasil simulasi pada kelompok peubah berkorelasi rendah dimana IPT3 terlihat memiliki tingkat validitas paling tinggi. Hal ini berarti keempat metode dapat dipertimbangkan untuk menggabungkan respon yang berkorelasi tinggi. Hasil selengkapnya dapat dilihat pada **Gambar 8**.



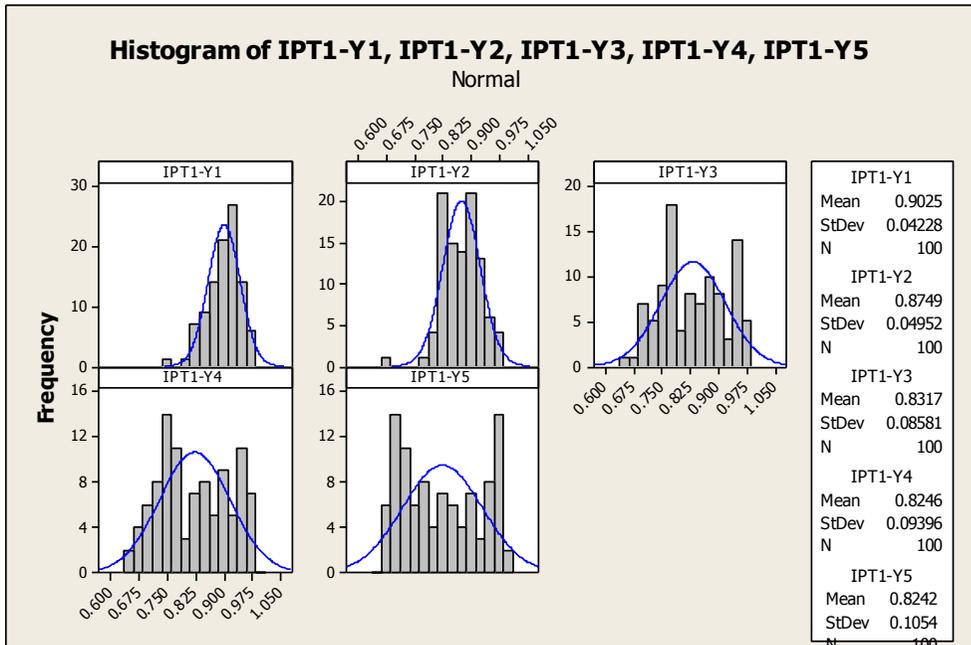
Gambar 8 Perbandingan tingkat validitas kelima metode penggabungan respon yang berkorelasi tinggi dari 100 kali simulasi

Namun demikian hasil ini juga harus diperkuat dengan melihat tingkat korelasi antara gabungan respon dengan setiap peubah asal. Gabungan respon yang baik diharapkan memiliki tingkat korelasi yang tinggi dengan seluruh peubah asal, jika hal ini terjadi berarti gabungan respon mampu menjadi wakil yang baik dalam menjelaskan perilaku seluruh peubah asal.

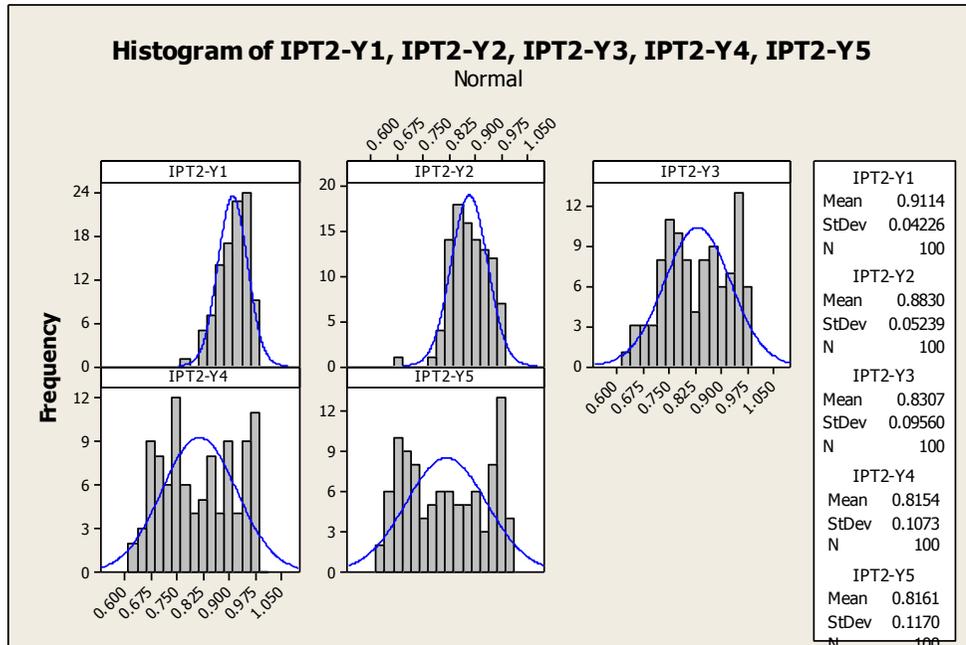
Model AMMI Kini dan Yang Akan Datang

Dari **Gambar 9** sampai **Gambar 13**, terlihat IPT1, IPT2, IPT3 dan IPT5 memiliki tingkat korelasi yang cukup besar dan merata dengan setiap peubah asal yaitu berkisar 0.8 sampai 0.9. Namun demikian jika dilihat dari kestabilan nilai korelasi yang dihasilkan, IPT3 tetap merupakan metode penggabungan yang terbaik. Hal ini terlihat dari nilai simpangan bakunya yang paling kecil yaitu berkisar antara 0.06 sampai 0.09. Kemudian secara berturutan diikuti oleh IPT5, IPT1 dan IPT2.

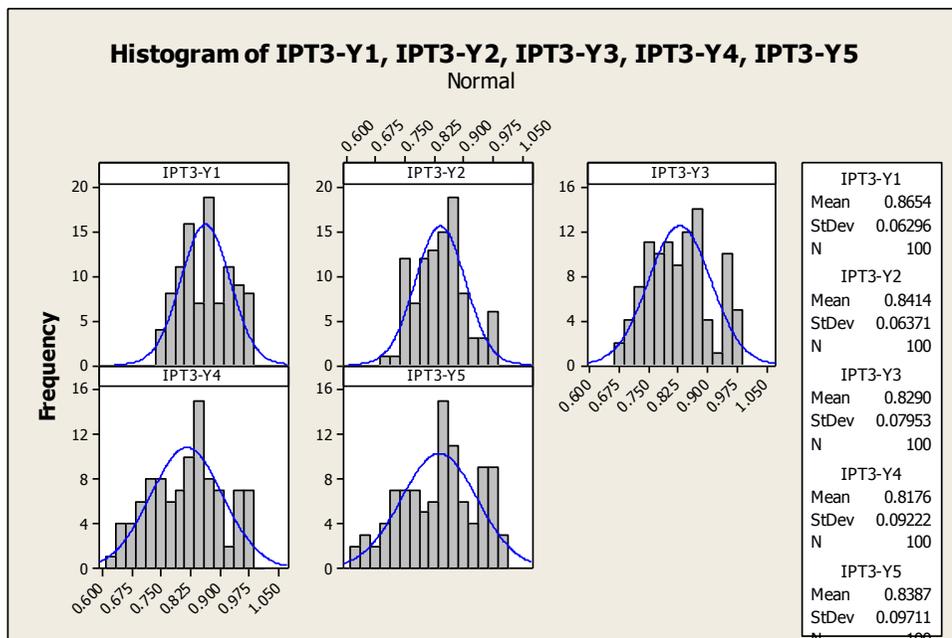
IPT4 bukan merupakan metode penggabungan respon yang dapat mewakili seluruh peubah asal dengan baik, hasil ini terlihat pada kelompok peubah berkorelasi rendah maupun tinggi.



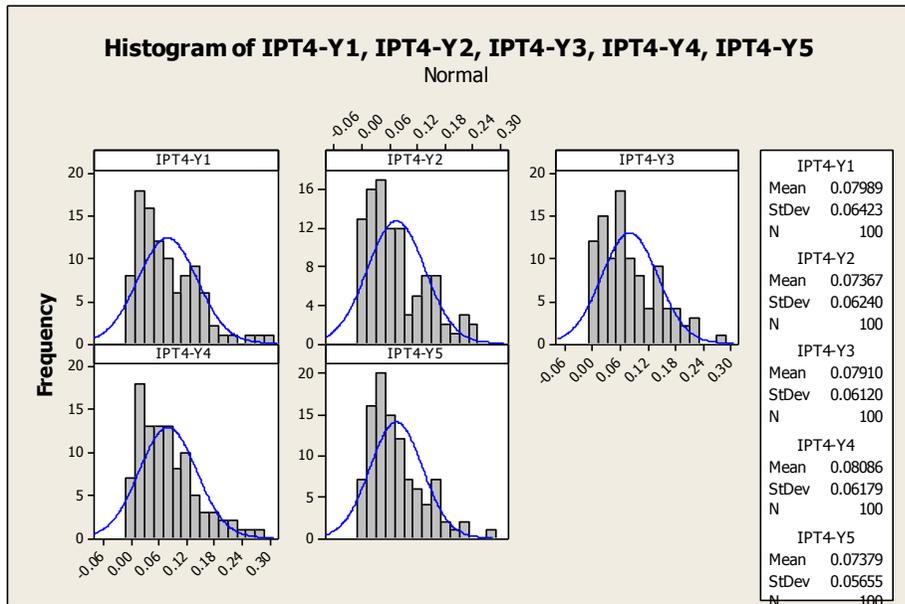
Gambar 9. Keragaan nilai mutlak korelasi antara gabungan respon dengan metode range equalization (IPT1) dengan seluruh peubah asal kelompok kedua dari 100 kali simulasi



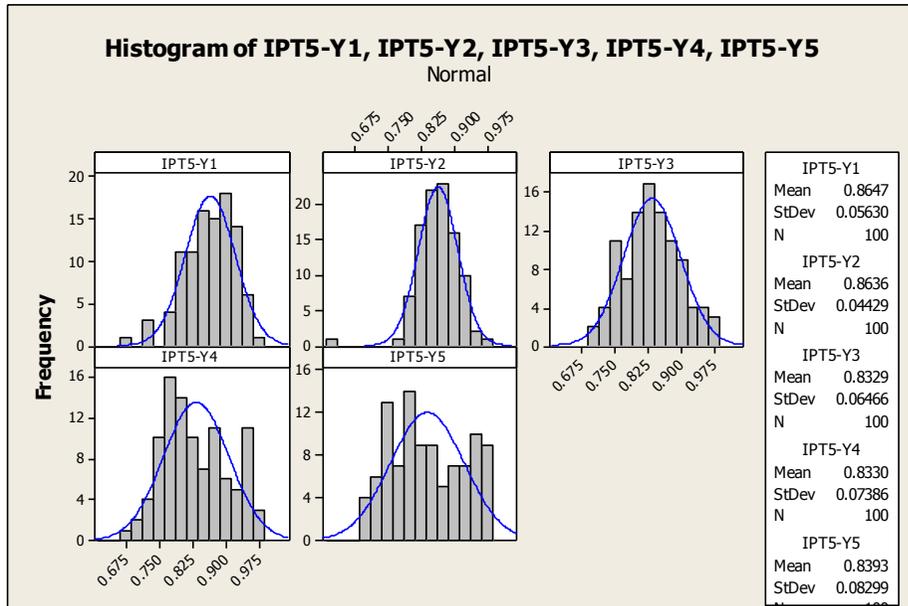
Gambar 10. Keragaan nilai mutlak korelasi antara gabungan respon dengan metode komponen utama pertama (IPT2) dengan seluruh peubah asal kelompok kedua dari 100 kali simulasi



Gambar 11. Keragaan nilai mutlak korelasi antara gabungan respon dengan metode pembobotan berdasarkan komponen utama (IPT3) dengan seluruh peubah asal kelompok kedua dari 100 kali simulasi



Gambar 12. Keragaan nilai mutlak korelasi antara gabungan respon dengan metode jarak Hotelling (IPT4) dengan seluruh peubah asal kelompok kedua dari 100 kali simulasi



Gambar 13. Keragaan nilai mutlak korelasi antara gabungan respon dengan metode division by mean (IPT5) dengan seluruh peubah asal kelompok kedua dari 100 kali simulasi

KESIMPULAN

Beberapa kesimpulan yang dapat ditarik antara lain:

1. Metode penggabungan respon dengan menggunakan sistem pembobotan berdasarkan komponen utama (IPT3) merupakan metode penggabungan respon yang terbaik, baik pada kelompok peubah berkorelasi rendah maupun tinggi.
2. Metode penggabungan respon yang juga memiliki hasil yang cukup baik adalah *divison by mean* (IPT5), *range equalization* (IPT1) dan skor komponen utama pertama (IPT2).
3. Metode jarak Hotelling (IPT4) kurang baik digunakan untuk menggabungkan respon karena tidak bisa mengikuti perilaku data asal, dimana korelasi minimum antara respon gabungan dengan peubah asal sangat rendah.

Ucapan Terimakasih

Tulisan ini bagian dari Hibah Penelitian Tim Pascasarjana yang didanai oleh Direktorat Jenderal Pendidikan Tinggi Nomor : 226/13.11/PL/2008 Tanggal : 02 April 2008

6

AMMI UNTUK MODEL CAMPURAN

Percobaan multilokasi merupakan teknik percobaan yang penting dalam bidang pemuliaan tanaman untuk mengetahui interaksi genotipe dengan lingkungan. Percobaan ini melibatkan dua faktor utama yaitu genotipe tanaman dan kondisi lingkungan. Salah satu model yang sering digunakan untuk melakukan analisis terhadap interaksi genotipe \times lingkungan (IGL) adalah model AMMI (Additive Main Effects and Multiplicative Interaction). Penguraian matriks interaksi menggunakan prosedur penguraian nilai singular (Singular Value Decomposition, disingkat SVD) dan penggunaan biplot merupakan kelebihan utama dari AMMI yang telah memberikan sumbangan besar bagi penelitian pemuliaan.

Selain SVD, prosedur pada analisis faktor juga dapat digunakan untuk menguraikan suatu matriks menjadi bentuk multiplikatif antara loading faktor dan skor faktor. Sehingga, analisis faktor dapat digunakan untuk menguraikan interaksi dalam konteks percobaan multilokasi. Sementara itu, jika

lingkungan yang dilibatkan dalam percobaan multilokasi merupakan sampel dari banyak lingkungan, maka pengaruh lingkungan dianggap sebagai pengaruh acak. Jika genotipe merupakan pengaruh tetap dan lingkungan acak, maka model linear yang dibentuk disebut model campuran. Prediksi dalam model campuran umumnya menggunakan metode prediksi takbias linier terbaik (Best Linear Unbiased Prediction, disingkat BLUP). Untuk mendapatkan penjelasan atas matriks interaksi yang bersifat acak, BLUP digabungkan dengan metode penguraian pengaruh interaksi. Karena itu, perlu dikaji metode yang lebih baik, juga perlu dikaji pengaruh penanganan model campuran terhadap model yang dihasilkan, jika dibandingkan dengan model tetap.

Sadiyah 2010 mememukakan bahwa, model AMMI dengan dekomposisi matriks interaksi dugaan menggunakan penguraian nilai singular memiliki kebaikan suai yang lebih tinggi jika dibandingkan dengan dekomposisi matriks interaksi dugaan menggunakan analisis faktor, pada kondisi banyaknya suku multiplikatif yang digunakan dalam model sama banyak. Hal ini berlaku baik pada kasus model tetap (fixed model) maupun model campuran (mixed model). Jika lingkungan yang dilibatkan dalam percobaan merupakan faktor acak, sehingga memiliki sebaran dengan keragaman tertentu, maka model campuran lebih tepat digunakan.

Masih menurut Sa'diyah 2010, penanganan keacakan faktor pada model AMMI campuran menunjukkan bahwa penggunaan BLUP - AMMI memberikan hasil pengepasan (fitting) dan klasifikasi yang dihasilkan sama dengan model tetap, namun model campuran memiliki struktur yang lebih sederhana. Model campuran lebih baik dibandingkan model tetap karena menghasilkan struktur model yang lebih sederhana dengan kemampuan pendugaan dan klasifikasi yang sama. Kesederhanaan struktur model merupakan salah satu pertimbangan dalam pemodelan.

Penanganan Model Campuran pada Percobaan Multilokasi Menggunakan BLUP

(Mixed Model in The Analysis of Multi-Environment Trial using BLUP)

Halimatus Sa'diyah¹⁾, A A Mattjik²⁾ dan I M Sumertajaya²⁾

¹⁾Mahasiswa Program Studi Statistika Sekolah Pasca Sarjana IPB, Dosen Fakultas
Pertanian Universitas Universitas

²⁾Dosen Departemen Statistika FMIPA IPB

Abstract

Multi-environment trials (MET) plays an important role in plant breeding research. Study concerning genotype-environment interaction (GEI) is needed in selection of stable genotype and estimates the yields of a genotype in particular environment. The MET includes two factors, genotype and environment. If environment assumed as random factor and genotype as fixed, then the model can be analyzed in term of mixed model which is estimating random effect using Best Linear Unbiased Prediction (BLUP). The AMMI (Additive Main effect and Multiplicative Interaction) analysis is one of statistical technique to analyze data from MET by combining analysis between analysis of variance and principal component analysis. One of the alternatives is factor analysis that decomposing the GEI as loading factors and factor. It is necessary to compare AMMI and Factor analytic (FA) models, and investigate the use of mixed models other than fixed. To get random environment data sets, simulation is conducted by random sub-setting from whole environment without replacement. So, here we have 30 mixed data sets. The result shows that AMMI has better goodness of fit than FA model when total variance explained by these model are the same, whether mixed model is better than fixed model in the context of its simpler model, and its similar goodness of fit and classification resulted.

Keywords: Mixed, BLUP, AMMI, Factor Analytic

PENDAHULUAN

Latar Belakang

Percobaan multilokasi merupakan teknik percobaan yang penting dalam bidang pemuliaan tanaman, terutama untuk mengetahui stabilitas genotipe. Percobaan ini melibatkan dua faktor utama yaitu genotipe tanaman dan kondisi lingkungan (lingkungan: tempat (*site*), perlakuan agronomis, musim). Umumnya, baik faktor genotipe maupun lingkungan diasumsikan sebagai faktor tetap, sehingga percobaan multilokasi dianalisis sebagai model tetap, dan pendugaan respon menggunakan metode kuadrat terkecil (MKT). Namun, jika lingkungan yang dilibatkan dalam percobaan multilokasi dianggap sebagai contoh acak dari populasi lingkungan yang mungkin, maka lingkungan dianggap sebagai faktor acak, sehingga model yang terbentuk disebut sebagai model campuran, dan metode pendugaannya menggunakan BLUP (*Best Linear Unbiased Prediction*).

Metode statistika yang biasa digunakan untuk analisis kestabilan genotipe pada percobaan multilokasi adalah AMMI (*Additive Main Effect and Multiplicative Interaction*). AMMI menguraikan matriks interaksi genotipe \times lingkungan (IGL) dengan prosedur penguraian nilai singular (PNS), kemudian mengkonstruksi biplot untuk memvisualisasi IGL tersebut. Biplot inilah yang menjadi kelebihan utama dari AMMI (Gauch, 2006). Selain PNS, prosedur pada analisis faktor juga dapat digunakan untuk menguraikan suatu matriks menjadi bentuk multiplikatif antara loading faktor dan skor faktor. Sehingga, analisis faktor dapat digunakan untuk menguraikan interaksi dalam konteks percobaan multilokasi.

Dalam tulisan ini dua hal yang menjadi fokus kajian pada model campuran bagi percobaan multilokasi adalah (i) pedugaan pengaruh interaksi menggunakan BLUP dan (ii) dekomposisi matriks interaksi menggunakan model Analisis Faktor. Secara teknis penelitian ini akan mengevaluasi kinerja keduanya, baik pada model tetap maupun model campuran dalam hal indeks stabilitas genotipe dan menduga respon.

Tujuan

Beberapa tujuan dari penelitian ini antara lain:

1. Membandingkan indeks stabilitas genotipe berdasarkan metode penguraian nilai singular (PNS) dan faktor analisis
2. Membandingkan akurasi pendugaan model AMMI dan model faktor analitik (FA) pada model tetap dan model campuran jika persentase keragaman yang dijelaskan oleh model sama

METODOLOGI

Data

Dalam penelitian ini menggunakan data riil dan data simulasi. Sumber data riil adalah percobaan multilokasi galur harapan padi sawah oleh konsorsium padi Nasional, melibatkan 14 genotipe dan 20 lingkungan. Respon yang diukur adalah hasil produksi padi (ton/ha). Percobaan menggunakan Rancangan Acak Kelompok pada petak berukuran 4 m x 5 m. Selanjutnya data 20 lingkungan ini dianggap data populasi. Sedangkan data simulasi dibangkitkan dengan memilih secara acak tanpa pemulihan 16 lingkungan dari 20 lingkungan pada data riil, diulang sebanyak 30 kali. Tahapan analisis data sebagai berikut:

- I. Melakukan analisis data menggunakan model tetap:
 1. Melakukan analisis ragam gabungan model tetap
 2. Menghitung rata-rata dari 3 ulangan untuk tiap genotipe di tiap lingkungan
 3. Menduga pengaruh utama genotipe dan lingkungan menggunakan

$$\text{Metode Kuadrat Terkecil (MKT), } \hat{\mu} = \frac{1}{ger} \sum_{i=1}^g \sum_{j=1}^e \sum_{k=1}^r y_{ijk} ; \hat{\tau}_i = \mu_i - \mu ;$$

$$\hat{\delta}_j = \mu_j - \mu ;$$

4. Mem-fit model menggunakan dugaan pengaruh utama

$$\tau \hat{\delta}_{ij} = \bar{y}_{ij} - \hat{\mu} - \hat{\tau}_i - \hat{\delta}_j$$
 untuk mendapatkan matriks interaksi dugaan
5. Menguraikan matriks interaksi dugaan yang dihasilkan dari tahap 4 menggunakan 2 metode:
 - a. Metode PNS, kemudian dicari skor komponen genotipe dan skor komponen lingkungan
 - b. Metode Faktor, kemudian dicari loading faktor (sebagai skor genotipe) dan skor faktor (sebagai skor lingkungan)
6. Membuat biplot AMMI dan Biplot Faktor berdasarkan skor komponen genotipe dan skor komponen lingkungan yang didapatkan dari tahap 5.
7. Menghitung indeks stabilitas AMMI (ISA) dan indeks stabilitas faktor (ISF) berdasarkan skor komponen genotipe dan skor komponen lingkungan yang didapatkan dari tahap 5.
8. Membuat model AMMI-5 dan model FA-5 dengan menggunakan 5 akar ciri terbesar yang dihasilkan tahap 5, dimana keragaman yang dijelaskan oleh model sekitar 75%.
9. Menduga respon daya hasil genotipe ke-i di lingkungan ke-j berdasarkan model yang dihasilkan dari tahap 6.

- II. Melakukan analisis data menggunakan model campuran
 1. Melakukan analisis ragam gabungan model campuran
 2. Menghitung penduga komponen ragam
 3. Menduga pengaruh genotipe, lingkungan dan interaksi menggunakan BLUP
 4. Menghitung rata-rata dari 3 ulangan untuk tiap genotipe di tiap lingkungan
 5. Ulangi langkah 1.4 – 1.9

Kriteria Evaluasi

Untuk mengukur kebaikan suatu model, pada penelitian ini digunakan kriteria sebagai berikut:

1. Akar dari kuadrat tengah galat (RMSE), dimana semakin kecil nilai RMSE maka semakin baik suatu model dalam menduga respon
2. Korelasi antara data asli dengan data dugaan, dimana semakin tinggi korelasinya, maka makin erat hubungan data asli dengan data dugaan. Hal ini dapat dijadikan indikasi tingkat kebaikan suai model, dimana makin tinggi korelasi, makin baik model tersebut
3. Korelasi antara indeks stabilitas data populasi 20 lingkungan dengan rata-rata indeks stabilitas model AMMI-m dan model FA-m yang dihasilkan dari 30 kali simulasi

Disamping ketiga kriteria evaluasi di atas, kesederhanaan struktur suatu model juga dipertimbangkan.

Percobaan Multilokasi

Percobaan multilokasi merupakan serangkaian percobaan yang serupa di beberapa lingkungan yang mempunyai rancangan percobaan dan perlakuan yang sama. Percobaan multilokasi dilakukan dengan melibatkan berbagai genotipe pada berbagai kondisi lingkungan. Percobaan ini memiliki tiga tujuan: a) menduga secara akurat dan memprediksi hasil berdasarkan data percobaan yang terbatas; b) menentukan stabilitas hasil dan pola respon genotipe terhadap lingkungan; c) menyediakan panduan yang *reliable* dalam memilih genotipe atau perlakuan agronomi terbaik untuk penanaman di masa yang akan datang dan di lokasi yang baru. Analisisnya sering disebut analisis ragam gabungan (*Composite Analysis of Variance*).

Model ANOVA dari percobaan multilokasi yang menggunakan rancangan acak kelompok (RAK) di tiap lingkungan adalah sebagai berikut:

$$y_{ijk} = \mu + \tau_i + \gamma_{k(j)} + \delta_j + (\tau\delta)_{ij} + \varepsilon_{ijk}, \quad i=1,\dots,g; j=1,\dots,e; k=1,\dots, r \quad (1)$$

Menurut Yang (2007), terdapat tiga bentuk yang mungkin dari persamaan (1) yaitu: (i) Model pengaruh tetap dengan semua komponen kecuali ε_{ijk} adalah tetap; (ii) Model pengaruh acak dengan semua pengaruh adalah acak kecuali μ ; (iii) Model pengaruh campuran, jika salah satu dari pengaruh genotipe ataupun lingkungan adalah tetap, dan yang lain adalah acak. Pada penelitian ini, pengaruh genotipe adalah tetap, sedangkan pengaruh lingkungan adalah acak, sehingga dari (1), μ dan τ_i adalah tetap sedangkan δ_j , $(\tau\delta)_{ij}$, $\gamma_{k(j)}$, dan ε_{ijk} saling bebas dan menyebar normal dengan nilai tengah nol dan ragam berturut-turut adalah σ_δ^2 , $\sigma_{\tau\delta}^2$, σ_γ^2 , dan σ_ε^2 .

Prediksi Takbias Linier Terbaik

Pada terminologi statistik, pendugaan suatu pengaruh acak disebut sebagai “prediksi” dan prosedur yang berhubungan dengan prediksi disebut BLUP (Searle *et al.*, 1992; Littell, *et al.*, 2006). Dengan merata-ratakan terhadap blok dalam lingkungan, persamaan (1) direduksi menjadi model yang sering digunakan untuk mendeskripsikan rata-rata sel, yaitu:

$$\bar{y}_{ij.} = \mu + \tau_i + \delta_j + (\tau\delta)_{ij} + \bar{\varepsilon}_{ij.} \quad (2)$$

dimana $\bar{\varepsilon}_{ij.} = \sum_{k=1}^r (\gamma_{k(j)} + \varepsilon_{ijk})/r$

Berdasarkan persamaan (2), diprediksi vektor dari daya hasil yang sebenarnya pada nilai \hat{y}_{ij} tertentu (\mathbf{w}_{ij}), dimana $\mathbf{w}_{ij} = \mu + \tau_i + \delta_j + (\tau\delta)_{ij}$. Dapat dilihat bahwa $\mathbf{E}(\mathbf{w}) = \mathbf{E}(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$. BLUP dari rata-rata sel ke-*ij* dihitung menggunakan metode yang dikembangkan oleh Yang (2007), yaitu:

$$BLUP(w_{ij}) = BLUE(\mu_i) + BLUP(\delta_j) + BLUP[(\tau\delta)_{ij}] \quad (3)$$

Dimana:

$$BLUE(\mu_i) = BLUE(\mu + \tau_i) = \text{nilai tengah genotipe ke-}i(\bar{y}_{i...}),$$

$$BLUP(\delta_j) = \frac{rg\sigma_\delta^2}{E(MS_\delta)}(\bar{y}_{.j.} - \bar{y}_{...}) \text{ dan}$$

$$BLUP[(\tau\delta)_{ij}] = \frac{r\sigma_{\tau\delta}^2}{E(MS_\delta)}(\bar{y}_{.j.} - \bar{y}_{...}) + \frac{r\sigma_{\tau\delta}^2}{E(MS_{\tau\delta})}(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})$$

Model AMMI (Additive Main Effect and Multiplicative Interactions)

Analisis AMMI adalah tehnik untuk analisis data percobaan dua arah, dimana pengaruh utama dimodelkan dengan bentuk aditif menggunakan prosedur ANOVA dan pengaruh interaksi dimodelkan dalam bentuk multiplikatif (Piepho, 1994; Sumertajaya, 2005). Langkah awal dari prinsip AMMI adalah mengepas (*fitting*) pengaruh utama aditif untuk genotipe dan lingkungan menggunakan prosedur ANOVA dan kemudian melakukan penguraian nilai singular (PNS) pada matriks sisaan kedalam m buah sumbu komponen utama interaksi (KUI).

Menurut Piepho (1994), model linear untuk AMMI adalah :

$$\bar{y}_{ij} = \mu + \tau_i + \delta_j + \sum_{s=1}^m \lambda_s \alpha_{is} l_{js} + \theta_{ij} \tag{4}$$

Misalkan T adalah matriks interaksi dengan ordo $g \times e$, dengan g adalah banyaknya genotipe dan e adalah banyaknya lingkungan, maka dekomposisi T dengan metode PNS adalah $T = A\Lambda L'$, dimana A adalah matriks ortogonal genotipe berukuran $g \times m$ yang kolom-kolomnya adalah vektor ciri dari TT' ; L adalah matriks ortogonal lingkungan berukuran $m \times e$ yang kolom-kolomnya adalah vektor ciri dari $T'T$; Λ adalah matriks diagonal berukuran $m \times m$ dengan elemen diagonal adalah nilai singular dari $T'T$.

Analisis Faktor

Analisis faktor merupakan perluasan dari analisis komponen utama. Analisis faktor digunakan untuk mengidentifikasi sejumlah kecil faktor yang dapat digunakan untuk menjelaskan sejumlah besar variabel yang saling berhubungan, dimana variabel-variabel dalam satu faktor berkorelasi tinggi, sedang korelasi dengan variabel pada faktor lain relatif rendah. Tiap-tiap kelompok dari variabel mewakili suatu konstruksi dasar yang disebut faktor. Untuk meningkatkan daya interpretasi faktor, harus dilakukan transformasi pada matriks loading (Dillon & Goldstein, 1984; Johnson & Winchern, 1998).

Menurut Smith, et al. (2007), model Faktor Analitik dengan m buah faktor (model FA-m) dapat dituliskan sbb:

$$\bar{y}_{ij} = \mu + \tau_i + \delta_j + \sum_{s=1}^m l_{is} f_{js} + \theta_{ij} \tag{5}$$

Asumsi yang digunakan: $E(\mathbf{F}) = \mathbf{0}$, $E(\underline{\epsilon}) = \mathbf{0}$, $Cov(\mathbf{F}) = E(\mathbf{FF}') = \mathbf{I}$, dan $Cov(\underline{\epsilon}) = E(\underline{\epsilon}\underline{\epsilon}') = \mathbf{\Psi} = \text{diag}(\Psi_1, \dots, \Psi_p)$, \mathbf{F} dan $\underline{\epsilon}$ saling bebas, $Cov(\underline{\epsilon}, \mathbf{F}) = E(\underline{\epsilon}\mathbf{F}') = \mathbf{0}$.

Hal yang membedakan persamaan untuk variabel g pada analisis faktor dengan regresi peubah ganda acak yang biasa adalah bahwa baik kovariat maupun koefisien regresi tidak diketahui, sehingga harus diduga dari data. Modelnya

adalah sebuah model multiplikatif dari koefisien lingkungan (skor faktor) dan koefisien genotipe (loading faktor).

Biplot dan Indeks Stabilitas

Biplot merupakan alat yang digunakan untuk menginterpretasi pola interaksi pada percobaan multilokasi (Mattjik & Sumertajaya, 2000). Biplot AMMI adalah plot antara skor KUI terbesar pertama (KUI1) dengan skor KUI terbesar kedua (KUI2) dari hasil PNS matriks interaksi. Skor KUI_s untuk genotipe ke-i adalah $\alpha_{is} \sqrt{\lambda_s}$ dan Skor KUI_s untuk lingkungan ke-j adalah $l_{js} \sqrt{\lambda_s}$ (i=1,...,g; j=1,...,e; s=1, 2).

Indeks stabilitas diperlukan untuk mempermudah melihat tingkat stabilitas suatu genotipe terhadap lingkungan. Indeks dibangun berdasarkan konsep jarak, sehingga semakin besar indeks suatu genotipe, maka semakin jauh jarak genotipe dari pusat sumbu koordinat, artinya semakin tidak stabil genotipe tersebut. Formula indeks stabilitas adalah sebagai berikut:

$$ISA = \sqrt{\left[\frac{\lambda_1^{1/2}}{\sum_{i=1}^m \lambda_i^{1/2}} (\text{SkorKUI}_1) \right]^2 + \dots + \left[\frac{\lambda_m^{1/2}}{\sum_{i=1}^m \lambda_i^{1/2}} (\text{SkorKUI}_m) \right]^2}$$

Dimana m adalah banyaknya komponen utama hasil dekomposisi matriks interaksi yang digunakan untuk menghitung indeks. Untuk biplot Faktor dan indeks stabilitas faktor (ISF) dibuat dengan cara yang sama dengan biplot AMMI dan indeks stabilitas AMMI (ISA), tetapi skor KUI digantikan dengan skor faktor. Skor faktor adalah hasil dekomposisi interaksi menggunakan metode faktor.

HASIL DAN PEMBAHASAN

Analisis Percobaan Multilokasi Padi Sawah

Analisis ragam gabungan model campuran menunjukkan bahwa pengaruh genotipe, lingkungan, interaksi genotipe × lingkungan (IGL) dan pengaruh kelompok nyata pada taraf $\alpha=0.05$, dimana masing-masing memiliki nilai-p < 0.0001. Sehingga dapat disimpulkan bahwa minimal ada sepasang genotipe memiliki nilai tengah yang berbeda, terdapat keragaman antar lingkungan yang berbeda, dan terdapat perubahan pengaruh faktor genotipe pada saat taraf faktor lingkungan berubah. Jika anova gabungan menggunakan model tetap, maka hasilnya juga sama, yaitu signifikan untuk semua sumber keragaman.

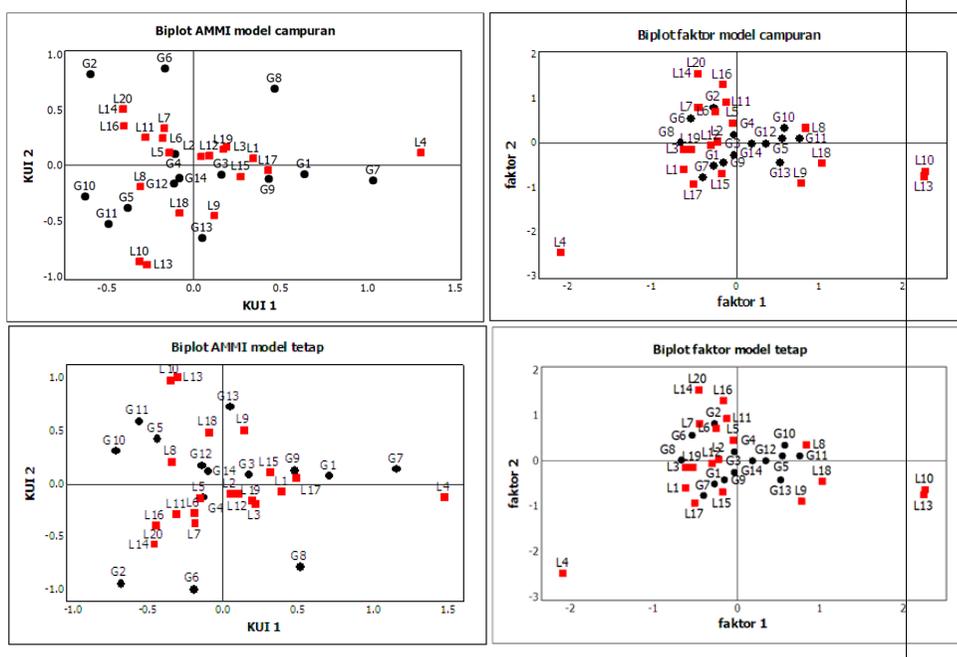
Model AMMI Kini dan Yang Akan Datang

Tabel 1 Lima akar ciri terbesar hasil dekomposisi menggunakan PNS dan Faktor

Akar ciri	PNS	Faktor	proporsi	Proporsi kumulatif
1	9.50606	05003	0.2309	0.2309
2	8.81120	0.4637	0.2140	0.4449
3	7.20428	0.3791	0.1750	0.6198
4	4.07317	0.2143	0.0989	0.7187
5	3.08845	0.1625	0.0750	0.7937

Hasil pendugaan nilai harapan kuadrat tengah dan komponen ragam menghasilkan semua komponen ragam positif, dengan nilai $p < 0.05$. Hal ini menunjukkan bahwa semua pengaruh acak memiliki keragaman yang nilainya berbeda nyata dengan nol. Penduga komponen ragam untuk lingkungan adalah yang terbesar, yaitu 1.8872. Penduga komponen ragam untuk pengaruh interaksi, blok, dan galat berturut-turut adalah 0.2125, 0.0742, dan 0.1756.

Pengaruh IGL yang nyata pada analisis ragam (ANOVA) menunjukkan bahwa dekomposisi matriks IGL akan bermakna. Hasil dekomposisi matriks interaksi dugaan menggunakan penguraian nilai singular (PNS) maupun faktor analisis menghasilkan 14 akar ciri positif tidak nol. Pada Tabel 1 dapat dilihat bahwa akar ciri yang dihasilkan dari dekomposisi interaksi dengan PNS dan faktor berbeda, namun memiliki persen keragaman yang sama.



Gambar 1 Biplot AMMI dan Faktor untuk model tetap dan model campuran Skor genotipe dan skor lingkungan yang diperoleh berdasarkan akar ciri-akar ciri tersebut divisualisasikan menggunakan biplot (Gambar 1). Biplot faktor dan

biplot AMMI dapat menerangkan keragaman sebesar 44.49%. Dari sisi genotipe, dapat dilihat bahwa G4, G3, G12 dan G14 terletak relatif dekat ke pusat sumbu, sehingga dapat dikatakan bahwa keempat genotipe tersebut relatif lebih stabil dibandingkan 10 genotipe yang lain, sedangkan G7 dan G2 relatif paling tidak stabil. Sedangkan dari sisi lingkungan, nampak bahwa L4 memiliki keragaman lingkungan yang terbesar, kemudian diikuti oleh L10, L13, L14, dan L20. Sebaliknya L5, L2 dan L12 memiliki keragaman yang relatif kecil, ini berarti semua genotipe padi dapat tumbuh dengan baik jika ditanam pada ketiga lingkungan tersebut. Jika dilihat dari sisi tempat tumbuh, G9 cocok dibudidayakan di L17, G13 di L9. Selain itu, L4 merupakan tempat yang sesuai untuk G7, L8 sesuai untuk G5, G10 dan G11.

Berdasarkan biplot faktor dapat dilihat bahwa L4, L10, dan L13 berada jauh dari titik pusat, sehingga memiliki panjang vektor dari (0,0) yang terpanjang. Hal ini berarti bahwa keragaman genetik di ketiga lingkungan tersebut yang dapat dijelaskan oleh faktor 1 dan faktor 2 adalah yang terbesar. Sedangkan genotipe yang terdekat dengan titik pusat sumbu adalah G3, G9, G14. Meskipun Biplot AMMI dan Biplot Faktor pada model yang sama terlihat berbeda, perbedaan tersebut disebabkan oleh perbedaan kuadran, sedangkan posisi relatif antar genotipe adalah sama, sehingga pola interaksi yang dijelaskan oleh keduanya tetap sama. Jika dilihat dari segi model tetap dan campuran maka nampak pola yang berbeda antara biplot AMMI model tetap dengan campuran, serta antara biplot faktor model tetap dengan campuran.

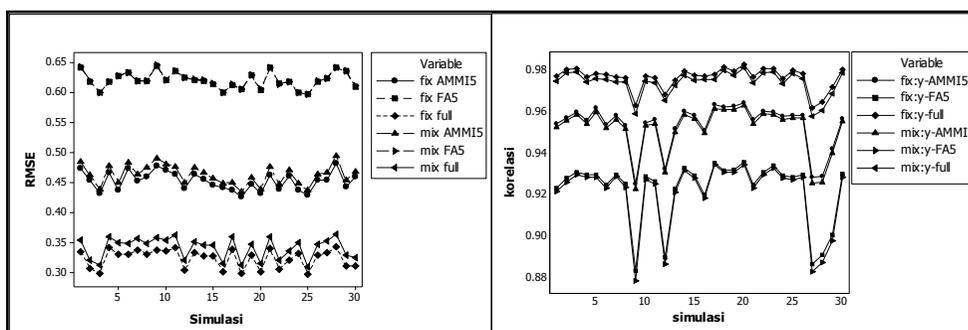
Kajian Simulasi

Simulasi data sebanyak 30 kali menghasilkan nilai korelasi dan rata-rata kuadrat tengah galat (RMSE) sebagaimana pada Gambar 2. Berdasarkan Gambar 2 dapat dilihat bahwa model penuh selalu lebih baik dibandingkan model tidak penuh yang hanya menggunakan sebagian komponen multiplikatif. Dari sudut pandang metode dekomposisi, yaitu jika dibandingkan model AMMI dengan model FA, jika banyaknya akar ciri yang digunakan sama, dalam hal ini digunakan sebanyak 5 komponen multiplikatif, maka korelasi model AMMI selalu lebih tinggi dari model FA, sedangkan RMSE model AMMI selalu lebih rendah nilainya dibandingkan model FA, baik pada model tetap maupun campuran. Sedangkan dari Gambar 3 dapat dilihat bahwa korelasi antara rata-rata indeks stabilitas AMMI (ISA) dengan indeks stabilitas populasi lebih tinggi dibandingkan indeks stabilitas faktor (ISF).

Model AMMI Kini dan Yang Akan Datang

Berdasarkan hal tersebut dapat diketahui bahwa akurasi pendugaan model AMMI lebih baik dibandingkan model FA, baik pada model tetap maupun model campuran jika persentase keragaman yang dijelaskan oleh model sama. Hal ini dikarenakan dalam model AMMI jika semua KUI digunakan, maka tidak ada sisaan matriks interaksi. Galat yang terjadi merupakan galat percobaan. Sedangkan pada model FA dimana dekomposisi matriks interaksi menggunakan metode faktor analisis, skor faktor adalah berupa dugaan dalam pendugaan tersebut terdapat komponen galat, sehingga meskipun seluruh faktor digunakan, masih menyisakan galat selain galat percobaan.

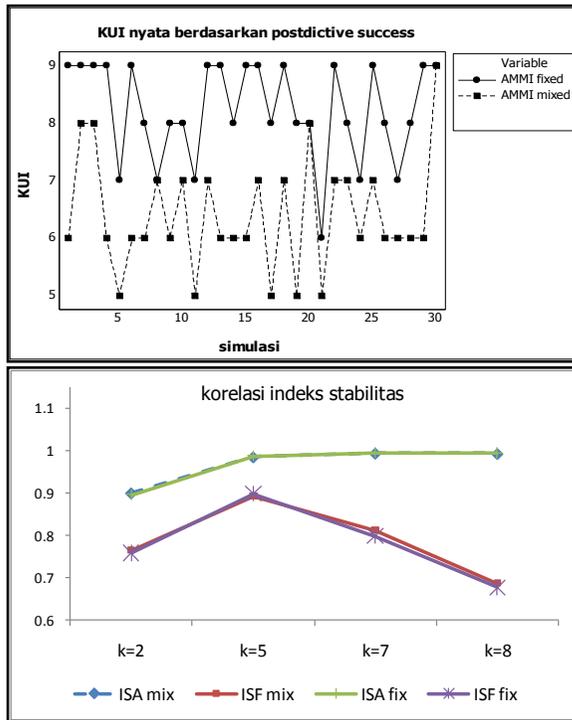
Jika dilihat dilihat dari sudut pandang model tetap dan model campuran, dalam kasus dimana lingkungan merupakan sampel dari populasi lingkungan sehingga lingkungan ditetapkan sebagai faktor acak, maka secara teori model campuran lebih tepat untuk digunakan. Hal ini dikarenakan orientasi dari suatu faktor acak adalah ragam, sehingga tidak bisa diperlakukan seperti faktor tetap yang berorientasi pada perbandingan nilai tengah. Sehingga prediksi pada model campuran harus diboboti terlebih dahulu dengan ragamnya.



Gambar 2 Korelasi dan RMSE model tetap dan campuran

Dari Gambar 2 dapat dilihat bahwa nilai RMSE dan korelasi untuk model tetap dan model campuran tidak terlalu berbeda. Bahkan pada model FA-5, nilai RMSE untuk model tetap dan campuran sangat mirip, sehingga digambarkan dalam satu garis. Hal ini dapat dijadikan indikasi bahwa model faktor lebih kekar terhadap pemilihan model, apakah akan digunakan model tetap ataupun campuran, dimana nilai RMSE-nya dapat dianggap sama.

Berdasarkan Gambar 3 dapat dilihat banyaknya komponen multiplikatif (KUI) hasil dari dekomposisi menggunakan PNS yang nyata pada uji F jika digunakan model tetap dan model campuran. Nampak bahwa banyaknya KUI nyata pada AMMI model tetap selalu sama atau lebih banyak dibandingkan AMMI model campuran.



Gambar 3 KUI nyata berdasarkan kriteria *postdictive success* dan korelasi indeks stabilitas

Hal ini berarti struktur model yang dihasilkan oleh AMMI model campuran lebih sederhana karena melibatkan lebih sedikit komponen multiplikatif. Korelasi indeks stabilitas untuk model tetap dan model campuran juga hampir sama, baik pada ISA maupun ISF. Sehingga berdasarkan struktur model dan kemampuan klasifikasi serta kemampuan pendugaan, maka model campuran lebih baik dibandingkan model tetap karena menghasilkan struktur model yang lebih sederhana dengan kemampuan pendugaan dan klasifikasi yang sama.

KESIMPULAN

Dalam penguraian matriks interaksi, jika dilihat dari kemampuan model untuk menduga respon, Model AMMI memiliki kebaikan suai yang lebih tinggi dibandingkan model faktor analitik jika persentase keragaman keduanya sama, baik dalam model campuran maupun model tetap.

Jika lingkungan yang dilibatkan dalam percobaan merupakan faktor acak, maka secara teori model campuran lebih tepat digunakan. Sedangkan berdasarkan

struktur model dan kemampuan klasifikasi serta kemampuan pendugaan, maka model campuran lebih baik dibandingkan model tetap karena menghasilkan struktur model yang lebih sederhana dengan kemampuan pendugaan dan klasifikasi yang sama.

DAFTAR PUSTAKA

- Dillon, W.R and M. Goldstein.** 1984. *Multivariate Analysis: Methods and Applications*. John Wiley & Sons, New York.
- Gauch, H.G. 2006. Statistical Analysis of Yield Trials by AMMI and GGE. *Crop Sci.* 46:1488-1500.
- Johnson, R.A and D.W. Winchern. 1998. *Applied Multivariate Statistical Analysis*. 4th ed. Prentice Hall, New Jersey.
- Littell, R.C, et al. 2006. *SAS for Mixed Model*. 2nd edition. <http://books.google.co.id/> , [5 Juli 2009].
- Mattjik, A.A dan I M. Sumertajaya. 2000. *Perancangan Percobaan dengan Aplikasi SAS dan Minitab*. IPB Press, Bogor.
- Piepho, H.P. 1994. Best Linear Unbiased Prediction (BLUP) for Regional Yield Trials: a Comparison to Additive Main Effect and Multiplicative Interaction (AMMI) Analysis. *Thoer. Appl. Genet.* 89: 647-654.
- Searle, S.R., G. Casella, Mc.Culloch, C.E. 1992. *Variance Components*. John Wiley and Sons, Canada.
- Smith, A., B.R. Cullis, R. Thompson. 2001. *Analysing Variety by Environment Data Using Multiplicative Mixed Models and Adjustment for Spatial Field Trends: Part 1: Theory*. In M.S. Kand (ed.) *Quantitative Genetics, Genomics, and Plant Breeding*. CABI Publishing, Oxford, UK.
- Sumertajaya I M. 2005. *Kajian Pengaruh Interblok dan Interaksi Pada Uji Multilokasi dan Respon Ganda [Disertasi]*. Bogor : Sekolah Pascasajana, Insitut Pertanian Bogor.
- Yang, R.C. 2007. Mixed Model Analysis of Crossover Genotype-Environment Interaction. *Crop Sci.* 47:1051-1062

7

PENGGUNAAN MODEL PERSAMAAN STRUKTURAL PADA MODEL AMMI

Seleksi genotipe unggulan seringkali sulit dilakukan karena nyatanya pengaruh interaksi genotipe \times lingkungan (IGL) pada percobaan multilokasi. Dibutuhkan pemahaman yang lebih dalam tentang IGL agar dapat membantu proses seleksi. Kajian tentang IGL telah banyak dilakukan diantaranya menggunakan metode *Additive Main Effect Multiplicative Interaction (AMMI)*. Metode AMMI dinilai berhasil dalam mengkaji struktur interaksi genotipe \times lingkungan dalam mengidentifikasi genotipe stabil dan spesifik lingkungan. Penguraian matrik interaksi dalam AMMI melalui *Singular Value Decomposition (SVD)* mampu memisahkan komponen multiplikatif dari galatnya (*noise*) (Gabriel, 1978 dalam Dhungana, 2004), sehingga penggunaan model AMMI mampu meningkatkan keakuratan dugaan respon interaksi genotipe \times lingkungan. Namun, model AMMI memiliki keterbatasan dalam hal ketidakmampuan

menjelaskan pengaruh dari kovariat genotipik dan lingkungan serta keterkaitan IGL beberapa karakteristik agronomi terhadap nyataanya interaksi genotipe \times lingkungan pada percobaan multilokasi.

Metode lain yang juga banyak digunakan mengkaji interaksi genotipe \times lingkungan diantaranya adalah metode *Factorial Regression* (FR) (Van Eeuwijk et al 1996 dalam Dhungana 2004) dan *Partial Least Square Regression* (PLSR) (Aastveit & Martens 1986 dalam Dhungana 2004). Kedua metode ini sukses dalam mengidentifikasi kovariat genotipik dan lingkungan yang paling berpengaruh terhadap nyataanya pengaruh interaksi genotipe \times lingkungan baik untuk IGL hasil ataupun IGL karakteristik agronomi yang lain. Namun kedua metode ini gagal dalam menjelaskan keterkaitan antara IGL karakteristik agronomi dengan IGL hasil sebagai suatu rangkaian proses fisiologis yang berkerja dalam sebuah sistem persamaan. Memahami keterkatian antara IGL karakteristik agronomi dengan IGL hasil tentunya merupakan bagian penting dalam pemuliaan karena hasil adalah akumulasi respon dari karakteristik agronomi terhadap kondisi lingkungan selama proses pertumbuhan dan perkembangan tanaman (Ivory 1989 dalam Noor et.al 2007).

Dhungana (2004), memperkenalkan penggabungan metode AMMI dengan model persamaan struktural (MPS) dalam menjelaskan interaksi genotipe \times lingkungan untuk hasil yang dikenal dengan MPS-AMMI. Melalui AMMI diperoleh bagian multiplikatif dari komponen interaksi dan mengeluarkan peubah galat (*noise*) sehingga pemodelan dengan MPS-AMMI menggunakan pola sesungguhnya dari interaksi geotipe \times lingkungan yang artinya model MPS-AMMI akan memberikan gambaran yang lebih tepat dalam menjelaskan nyataanya efek interaksi genotipe \times lingkungan untuk hasil. Melalui MPS-AMMI dapat dilakukan pemodelan IGL dengan memperhatikan proses fisiologis pertumbuhan dan perkembangan genotipe yang menjelaskan bagaimana keterkaitan IGL karakteristik agronomi dan bagaimana pengaruhnya terhadap IGL hasil dengan memperhatikan kekeliruan pengukuran dan memberikan informasi kecocokan model (*goodness of fit*) sebagai indikator kemampuan model dalam menjelaskan keragaman data. MPS-AMMI juga mampu menjelaskan bagaimana pengaruh kombinasi kovariate genotipe dengan lingkungan terhadap interaksi genotipe \times lingkungan untuk karakteristik agronomi dan hasil.

Kajian MPS-AMMI dapat digunakan untuk mengidentifikasi pada kondisi lingkungan dan karakteristik seperti apa genotipe-genotipe akan memberikan hasil yang lebih baik. Dengan kata lain, kajian ini memberikan informasi awal kepada pemulia tanaman untuk lebih fokus pada karakteristik genotipe dan faktor lingkungan yang paling berperan dalam peningkatan hasil. Tahapan paling penting dalam memodelan struktur interaksi dengan MPS-AMMI adalah mendapatkan peubah laten IGL karakteristik agronomi.

Pendugaan parameter model persamaan structural bergantung pada kesyaran asumsi distribusi dari data. Jika asumsi distribusi normal multivariate, pendugaan

parameter model dapat menggunakan metode maximum likelihood namun jika asumsi ini terlanggar maka pendugaan yang dapat digunakan adalah robust maximum likelihood, Weighted Least Square. Pendekatan terbaru dalam pendugaan parameter model persamaan structural adalah menggunakan metode Partial Least Square. Metode ini tidak membutuhkan ukuran sampel besar seperti halnya dengan menggunakan metode maximum likelihood ataupun metode-metode pendugaan lain dalam model persamaan structural yang berbasis struktur kovarians. Sedangkan Partial Least Square adalah metode pendugaan parameter persamaan structural yang berbasis component varians. Software yang dapat digunakan adalah AMOS, LISREL, Visual VLS dan lain-lainnya.

Berikut ini disajikan secara lengkap dua artikel tentang penggunaan *SEM* dan *Partial Least Square path modeling* pada model AMMI dan

Halaman ini sengaja dikosongkan

7.1 ANALISIS INTERAKSI GENOTIPE X LINGKUNGAN MENGGUNAKAN STRUCTURAL EQUATION MODELING (I M Sumertajaya, A A Matjjik, I G N M Jaya, Journal Pythagoras Vol. 4, No. 1, hal. 15-32)

Analisis Interaksi Genotip × Lingkungan Menggunakan Structural Equation Modeling

I Made Sumertajaya¹, Ahmad Ansori Matjjik² &
I Gede Nyoman Mindra Jaya³

¹) Departemen Statistika, FMIPA Institut Pertanian Bogor

²) Guru Besar Statistika, FMIPA Institut Pertanian Bogor

³) Mahasiswa S2 Statistika, Sekolah Pascasarjana Institut Pertanian Bogor

Abstract

Additive Main Effect and Multiplicative Model (AMMI Model) nowadays is used to asses in plant breeding, especially to asses the Genotype × Environment Interaction (GEI) on multi-environment trial. The presence of genotype × environment interaction (GEI) creates difficulties in modeling complex trait that involve sequence biological process. Coupling Structural equation modeling with AMMI was developed to analyzed genotype × environment interaction (GEI). Structural equation modeling allows us to account for underlying sequential process in plant development by incorporating intermediate variables associated with those processes in the model. With this method we can incorporating genotypic and environmental covariate in the model and explain how those covariates influence grain yield. SEM-AMMI useful when both environments and genotype are fixed and the purpose of the multi-environment trials (MET) is to assess the combined effect genotypic and environmental covariate on yield and yield components

Keywords : AMMI Model, Structural equation modeling

LATAR BELAKANG

Percobaan multilokasi merupakan percobaan yang sering dilakukan dalam penelitian pemuliaan tanaman untuk mengkaji interaksi genotip × lingkungan (*Genotype × Environmental Interaction = GEI*). *GEI* dapat dinyatakan sebagai perubahan keragaman dari beberapa genotip pada beberapa lingkungan berbeda. Kajian ini penting dalam pemuliaan tanaman karena hasilnya dapat digunakan untuk menduga dan menyeleksi genotip-genotip yang berpenampilan stabil (*stability of genotype*) pada berbagai lingkungan berbeda atau beradaptasi pada suatu lingkungan spesifik (*adaptation of genotypes to specific environmental*).

Daya hasil merupakan karakteristik yang sangat kompleks yang merupakan akumulasi respon terhadap kondisi lingkungan selama tahapan perkembangan dari sejumlah komponen daya hasil yang saling terkait sehingga dibutuhkan pemahaman aspek fisiologis yang mendasari *GEI* dalam upaya menemukan varietas unggulan. Melalui pemahaman konsep ini, maka dalam upaya menemukan varietas unggulan harus dimulai dengan memperhatikan faktor-faktor *genotypic* (respon dari genotip) dan lingkungan sebagai salah satu elemen yang bertanggung jawab atas signifikansi *GEI* pada daya hasil dengan mempertimbangkannya sebagai penyebab *GEI* selama proses perkembangan secara bertahap dari sejumlah komponen daya hasil yang saling terkait, dan menghubungkannya dengan daya hasil.

Beberapa metode statistik telah dikembangkan untuk mengkaji *GEI*, namun sedikit sekali metode yang melibatkan proses fisiologis, faktor-faktor *genotypic* dan lingkungan dalam menjelaskan *GEI* karena umumnya dalam pemuliaan tanaman daya hasil yang menjadi pusat perhatian dari para pemulia. Namun, jika hanya memperhatikan daya hasil dalam analisis *GEI* tentunya tidak cukup dalam pemilihan varietas unggulan yang stabil pada berbagai lokasi yang berbeda, sehingga dibutuhkan suatu metode yang efektif untuk menjelaskan efek dari faktor-faktor *genotypic* dan lingkungan terhadap *GEI* dalam mempengaruhi daya hasil.

Tai (1979) menyatakan sangatlah penting untuk mempertimbangkan keterkaitan antara proses fisiologis ketika memodelkan *GEI*. Hasil observasi menunjukkan bahwa hasil *GEI* untuk daya hasil meningkat ketika proses fisiologis untuk genotip yang berbeda menunjukkan respon yang berbeda terhadap faktor-faktor lingkungan. Tai menggunakan analisis jalur untuk menjelaskan keterkaitan proses fisiologis terhadap *GEI* untuk daya hasil. Pendekatan ini memiliki keterbatasan karena tidak mampu menjelaskan pengaruh dari variabel eksogenus seperti iklim dan faktor genetik terhadap *GEI* untuk daya hasil ataupun

GEI untuk komponen daya hasil. Sehingga dibutuhkan suatu metode yang memungkinkan memberikan penjelasan pengaruh dari faktor-faktor *genotypic* dan lingkungan terhadap signifikansi *GEI* untuk daya hasil dan *GEI* untuk komponen daya hasil.

Additive main effects and multiplicative interaction model (AMMI model) merupakan suatu metode multivariat yang relatif baru digunakan akhir-akhir ini dalam penelitian pemuliaan tanaman untuk mengkaji *GEI* pada suatu percobaan multilokasi. Gauch dan Zobel (1990) mengemukakan bahwa model *AMMI* merupakan suatu model gabungan dari pengaruh aditif pada analisis ragam dan pengaruh multiplikatif pada analisis komponen utama. Melalui *AMMI*, matriks residual sebagai penyimpangan dari model aditif didekomposisi dengan menggunakan *singular value decomposition (SVD)* untuk mendapatkan bagian multiplikatif dari model (Gabriel, 1978). *Singular value decomposition* memusatkan pola utama dari variasi residual kedalam sedikit komponen utama dan sisanya adalah *noise*. Dengan cara ini, bagian multiplikatif pada *AMMI* dapat mempartisi data ke dalam pola model interaksi penuh dengan mempertimbangkan sedikit komponen utama dan membuang *noise* untuk ketepatan prediksi (Gauch and Zobel 1996). Romogosta *et al.* (1993) menemukan model *AMMI* dengan dua komponen memiliki kemampuan paling akurat dalam melakukan prediksi dan memiliki kemudahan dalam interpretasi.

Metode *Factorial Regression (FR)* dan *Partial Least Square (PLS)* telah digunakan untuk memodelkan *GEI* dengan memasukkan kovariat *genotypic* dan lingkungan, namun gagal menjelaskan keterkaitan antar variabel penelitian (Dugana, 2004). Metode baru dikemukakan oleh Dugana sebagai perbaikan dari metode-metode yang telah ada yaitu menggunakan *Structural equation modeling (SEM)*. *SEM* merupakan sebuah pendekatan dalam melakukan analisis secara simultan untuk sebuah sistem persamaan yang mana setiap persamaan menjelaskan hubungan sebab – akibat antara variabel yang dimasukkan dalam sistem. *SEM* dapat digunakan untuk memodelkan keterkaitan antara komponen daya hasil dan menghubungkannya dengan daya hasil. *SEM* juga dapat menguji model dengan beberapa variabel dependen, memasukkan kekeliruan pengukuran dan mampu melakukan pengujian koefisien *path* secara parsial.

BAHAN DAN METODE

Data Eksperimen

Data yang digunakan dalam penelitian ini adalah data pemuliaan Jagung Hibrida yang dilakukan dari tanggal 23 Juli 2006 sampai 10 April 2007 yaitu pada musim

Model AMMI Kini dan Yang Akan Datang

hujan dan kemarau. Percobaan ini menggunakan 7 galur Jagung Hibrida Harapan dan 5 galur Jagung Hibrida Komersial yang ditanam pada 18 kondisi lokasi tersebar di 6 Propinsi.

Tabel 1. Deskripsi Lokasi Penelitian

No	Propinsi	Kecamatan	Desa	Elevasi (m)	Musim 2006/2007	
					Kemarau	Hujan
1	Jawa Barat	Cigombong	Ciburuy	521		L1
2	Jawa Barat	Bogor Barat	Pabuaran	260	L11	
3	Jawa Tengah	Banyodono	Nagru Aru	196	L12	
4	Jawa Tengah	Banyodono	Ketaon	190		L2
5	Jawa Tengah	Gemblengan	Kalikotes	190	L13	
6	Jawa Timur	Ambulu	Pontang	10	L17	
7	Jawa Timur	Kedung Mulyo	Brodot	60		L7
8	Jawa Timur	Tajinan	Jambu Timur	465	L18	
9	Jawa Timur	Tumpang	Wringinsongo	540		L8
10	Lampung	Ratu Nuban	Sido waras	35		L6
11	Lampung	Metro Timur	Yoso Mulyo	50	L16	L5
12	Sulawesi Selatan	Barru	Kemiri	45		L3
13	Sulawesi Selatan	Moncongloe	Moncongloe Bulu	17		L4
14	Sumatera Utara	Namo Rambe	Kuta Tengah	95	L14	L9
15	Sumatera Utara	Binjai	Sambirejo	35	L15	
16	Sumatera Utara	Sei Rampah	Cempedak Lobang	65		L10

Tabel 2. Jenis Galur

No.	Galur/varietas	Asal	Kelompok
A	BIO 9900	Bioseed	Hybrid
B	BIO 1263	Bioseed	Hybrid
C	BIO 1169	Bioseed	Hybrid
D	BC 42521	Bioseed	Hybrid
E	BC 42683	Bioseed	Hybrid
F	BC 41399	Bioseed	Hybrid
G	BC 2630	Bioseed	Hybrid
H	BC 42882 -A	Bioseed	Hybrid
I	BIO 9899	Bioseed	Hybrid
J	BISI - 2	PT. BISI	Hybrid
K	P - 12	PT. Dupont	Hybrid
L	C 7	PT. Monsanto	Hybrid

Dalam penelitian ini, variabel-variabel yang akan diambil adalah :

Tabel 3. Variabel yang Diamati

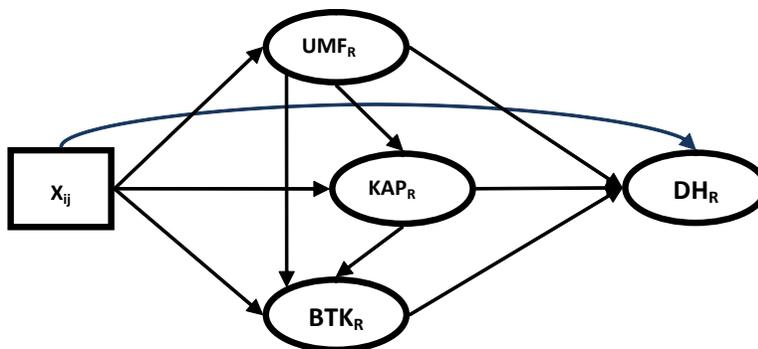
Variabel Yang Diamati	Satuan
Umur Masak Fisiologis (UMF)	Hari
Kadar Air saat panen (KAP)	MI
Berat Tongkol (BTK)	Gram/Tkl
Daya Hasil (DH)	Ton/Ha
Tinggi Lokasi (TL)	M
Musim (M)	0=Kemarau; 1=Hujan

Metode Analisis

Beberapa tahapan analisis dilakukan dalam penelitian ini meliputi analisis ragam dengan pendekatan ANOVA klasik, melakukan pemodelan pengaruh interaksi dengan metode AMMI. Melalui metode AMMI diperoleh variabel latent endogen yang akan digunakan dalam analisis *Structural equation modeling*. SEM digunakan untuk melihat bagaimana pengaruh dari komponen daya hasil terhadap daya hasil (DH_R , UMF_R , KAP_R , dan BTK_R), dan interaksi kovariat galur dengan kovariat faktor-faktor lokasi (X_{ij}) terhadap komponen daya hasil dan daya hasil sendiri.

HIPOTESIS PENELITIAN

Hipotesis yang diajukan dalam penelitian ini adalah bahwa daya hasil dapat dijelaskan oleh komponen daya hasil yaitu umur masak fisiologis, kadar air saat panen dan berat tongkol.



Gambar 1. Hipotesis Penelitian

Kemudian masih-masing komponen daya hasil dan daya hasil itu sendiri dapat dijelaskan oleh interaksi faktor-faktor lingkungan dan genetik. Hipotesis penelitian di atas dapat digambarkan dalam diagram alur gambar 1.

TAHAPAN ANALISIS

Tahap Pertama

Pada tahap pertama, data dianalisis dengan analisis ragam klasik untuk memperoleh gambaran tentang pengaruh utama dan pengaruh interaksi dari galur dan lokasi terhadap komponen daya hasil (UMF, KAP, BT) daya hasil (DH), dengan model RAL nya adalah :

$$y_{ger} = \mu + \alpha_g + \beta_e + \gamma_{ge} + \varepsilon_{ger} \quad (1)$$

Jika dalam analisis ini ditemukan adanya pengaruh interaksi antara lingkungan dengan galur, maka analisis akan diteruskan pada tahap ke dua.

Tahap Kedua

Pemodelan Analisis *AMMI*

Pada tahap kedua dilakukan analisis *AMMI* dengan tujuan menjelaskan interaksi perlakuan dengan lokasi, dan mendapatkan nilai komponen *AMMI* yang akan digunakan sebagai skor faktor endogen untuk analisis *SEM*. Untuk mendapatkan skor faktor endogen, dilakukan dengan menggunakan pendekatan Penguraian nilai singular (*SVD*).

Penguraian Nilai Singular

Penguraian Nilai Singular (*Singular Value Decomposition*) untuk matriks pengaruh interaksi **Z** sebagaimana dikemukakan oleh Greenacre (1984) adalah memodelkan matriks tersebut sebagai berikut:

$$\mathbf{Z} = \mathbf{U} \mathbf{L} \mathbf{A}' \quad (2)$$

Tujuan dari penguraian nilai singular ini adalah untuk mendapatkan faktor skor genotypic dan faktor skor lingkungan.

Penentuan Banyak Komponen *AMMI*

Penentuan banyak komponen *AMMI* digunakan metode *Postdictive success* berhubungan dengan kemampuan suatu model yang tereduksi untuk menduga data yang digunakan dalam membangun model tersebut. Salah satu penentuan banyaknya komponen berdasarkan *Postdictive success* adalah berdasarkan banyaknya sumbu tersebut yang nyata pada uji F analisis ragam.

Intepretasi Model AMMI

Alat yang digunakan untuk menginterpretasi hasil dari metode AMMI adalah biplot. Pada penelitian pemuliaan tanaman dikenal dengan istilah GGE biplot (*Genotype vs Genotype×Environment biplot*). Pada dasarnya metode ini merupakan upaya untuk memberikan peragaan grafik dari suatu matriks dalam suatu plot dengan menumpangtindihkan vektor-vektor dalam ruang berdimensi dua. Vektor-vektor yang dimaksud yaitu vektor yang mewakili nilai skor komponen lingkungan.

Tahap Ketiga

Memodelkan Pengaruh Komponen Daya Hasil Dan Kovariate Lingkungan × Genetik Terhadap Daya Hasil dengan Analisis SEM.

Karena tujuan kita adalah memodelkan daya hasil *GEI*, nilai daya hasil dan komponen daya hasil reseidual (DH_R , UMF_R , KAP_R , dan BT_R) digunakan sebagai variabel endogen (Y). Residual diperoleh dari hasil pengurangan nilai observasi dengan nilai dugaan efek utama galur dan lokasi. Variabel eksogen (X_{ij}) merupakan hasil perkalian antara kovariat galur ke- i dan kovariat lokasi ke- j . Karena daya hasil dan komponen daya hasil merupakan nilai residual, maka variabel eskogen (X) juga harus disesuaikan terhadap efek utama lokasi dan genotip dengan mengalikan nilai X dengan $(I-Pz)$ dimana Z adalah matriks rancangan dari efek utama galur dan lokasi, dan $Pz=Z(Z'Z)^{-1}Z'$ (Dhungana, 2004). Kita mengasumsikan bahwa variabel X diukur tanpa kesalahan pengukuran.

SEM-AMMI adalah model *structural equation modeling* dengan komponen multiplikatif *GEI* sebagai varaibel laten). Msalkan $Y_1, Y_2, Y_3,$ dan Y_4 masing-masing adalah matriks dugaan residual $DH_R, UMF_R, KAP_R,$ dan BT_R dengan ordo masing-masin $n \times p$ diman n adalah banyaknya genotip dan p adalah banyaknya lokasi. Setiap matriks residual dapat didefinisikan menggunakan matriks *singular value decomposition (SVD)*.

$$Y_i = U_i \Omega_i V_i + e_i \tag{3}$$

dengan :

- Y_i : $n \times p$ matriks residual variabel ke- i
- U_i : $n \times k$ matriks eigen vektor k pertama yang terkait dengan n genotipe
- Ω_i : $k \times k$ matriks diagonal nilai eigen k pertama dari nilai singular Y_i dengan elemen diagonal λ_i dimana $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$
- V_i : $k \times p$ matriks eigen vektor k pertama yang terkait dengan p lokasi

Kita asumsikan bahwa $U_i \Omega_i V_i$ adalah nilai *GEI* sebenarnya dari variabel ke- i dengan k komponen pertama ditentukan berdasarkan pada proporsi jumlah

kuadrat GEI yang ingin dijelaskan. Matriks Y_i dalam persamaan (3) dikonversi ke vektor kolom dengan menggunakan operator **vec** dan produk **konecker** (Harville, 1997) :

$$\mathbf{Vec}(Y_i) = (\mathbf{V}'_i \otimes \mathbf{U}_i) \mathbf{vec}(\Omega_i) + \mathbf{vec}(e_i) \quad (4)$$

dengan :

$\mathbf{vec}(Y_i)$: kolom Y_i disusun menjadi vektor kolom $np \times 1$

Misalkan

$$\eta_i = (\mathbf{V}'_i \otimes \mathbf{U}_i) \mathbf{vec}(\Omega_i) \quad (5)$$

Sehingga observasi dari genotip pada lokasi untuk variabel ke- i dituliskan sebagai berikut:

$$y_i = \eta_i + \varepsilon_i \quad (6)$$

Selanjutnya untuk keempat variabel, model pengukuran dari y dapat dituliskan sebagai berikut :

$$\mathbf{y} = \boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad (7)$$

dengan $\mathbf{y} = (y_1 \ y_2 \ y_3 \ y_4)'$, $\boldsymbol{\eta} = (\eta_1 \ \eta_2 \ \eta_3 \ \eta_4)'$, vector residual $\boldsymbol{\varepsilon} = (\varepsilon_1 \ \varepsilon_2 \ \varepsilon_3 \ \varepsilon_4)'$ dan $E(\boldsymbol{\varepsilon}) = 0$, $E(\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}') = \boldsymbol{\Theta}_\varepsilon$. Diasumsikan bahwa variabel eksogen (X) diukur tanpa kesalahan pengukuran.

Model strukturalnya dapat dituliskan sebagai berikut :

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\mathbf{X} + \boldsymbol{\zeta} \quad (8)$$

dengan \mathbf{X} : vektor ($s \times 1$) eksogenus; \mathbf{B} : matrik (4×4) koefisien yang menunjukkan hubungan antara variabel endogenus ($\boldsymbol{\eta}$); $\boldsymbol{\Gamma}$: Matriks ($4 \times s$) koefisien hubungan antara endogenus ($\boldsymbol{\eta}$) dengan eksogenus (X); $\boldsymbol{\zeta}$: vektor kolom (4×1) vektor kekeliruan yang terkait dengan variabel endogenus ($\boldsymbol{\eta}$).

Struktur Kovarians dan Pendugaan Parameter

Konsep pendugaan parameter dalam *SEM* adalah meminimumkan peredan antara kovarians observasi Σ dengan kovarians model $\Sigma(\theta)$ (Bollen, 1989; Johnson and Wichern, 1996).

Misalkan Σ_{yy} (4×4), Σ_{xx} ($s \times s$), Σ_{yx} ($4 \times s$) masing-masing adalah matriks kovarians dari 4 variabel observasi endogenus (Y), s variabel Eksogen (X), dan perkalian silang matriks kovarians (Y, X). Dan Σ adalah matriks gabungan dengan ordo $(4+s) \times (4+s)$ sebagai berikut :

$$\Sigma = \begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{yx} & \Sigma_{xx} \end{bmatrix}$$

$\Sigma(\theta)$ adalah matriks kovarians Y dan X yang merupakan fungsi dari vektor parameter (θ). Dengan menggunakan persamaan (7), (8) dan bentuk tereduksi dari persamaan (9) :

$$\eta = (I-B)^{-1}(\Gamma X + \zeta) \tag{9}$$

Sehingga partisi $\Sigma(\theta)$ yang sesuai dengan Σ adalah :

$$\Sigma_{yy}(\theta) = E(\mathbf{Y}\mathbf{Y}') = (I-B)^{-1}(\Gamma \Sigma_{xx} \Gamma' + \psi) [(I-B)^{-1}]^{-1} + \Theta_\varepsilon \tag{10}$$

$$\Sigma_{yx}(\theta) = E(\mathbf{Y}\mathbf{X}') = (I-B)^{-1}\Gamma \Sigma_{xx} \tag{11}$$

$$\Sigma_{xx}(\theta) = E(\mathbf{X}\mathbf{X}') = \Sigma_{xx} \tag{12}$$

Sehingga dapat dituliskan matriks $\Sigma(\theta)$ lengkapnya adalah :

$$\Sigma(\theta) = \begin{bmatrix} (I-B)^{-1}(\Gamma \Sigma_{xx} \Gamma' + \psi) [(I-B)^{-1}]^{-1} + \Theta_\varepsilon & (I-B)^{-1}\Gamma \Sigma_{xx} \\ \Sigma_{xx} \Gamma' [(I-B)^{-1}]' & \Sigma_{xx} \end{bmatrix} \tag{13}$$

Pendugaan Parameter :

Beberapa metode pendugaan parameter dikenal dalam SEM diantaranya adalah

Pendugaan Maksimum Likelihood

$$F_{ML} = \log |\Sigma| + tr \left[\Sigma^{-1} S \right] - \log |S| - p + q \tag{14}$$

Uji Kesesuaian Model

Uji Kebaikan Suai Khi-Kuadrat

Ukuran kebaikan ini pada dasarnya merupakan pengujian seberapa dekat matriks hasil dugaan dengan matriks data asal dengan menggunakan uji khi-kuadrat. Hipotesis yang diuji adalah

$$H_0: \Sigma = \Sigma(\theta) \quad \text{lawan} \quad H_1: \Sigma \neq \Sigma(\theta)$$

Jika H_0 diterima pada taraf signifikan tertentu, maka dapat diambil kesimpulan bahwa model diterima. Statistik untuk menguji hipotesis tersebut adalah:

$$\chi^2 = (n-1) F \tag{15}$$

Statistik tersebut mendekati distribusi khi-kuadrat. Jika nilai χ^2 lebih besar dari nilai kritis khi-kuadrat dengan taraf signifikansi $\chi^2(df, \alpha)$ maka H_0 ditolak.

SOFTWARE

Untuk mempermudah perhitungan dalam penelitian ini, penulis menggunakan beberapa software yaitu Excel 2007, SAS 9.1 dan AMOS 7.0

HASIL DAN PEMBAHASAN

Melalui uji Levene's diperoleh nilai statistik (t-student) sebesar 4.84 dengan nilai peluang nyata sebesar 0.000. Hasil pengujian ini menunjukkan bahwa data daya hasil jagung tidak homogen untuk setiap lokasi. Setelah dilakukan eksplorasi data ditemukan bahwa data daya hasil jagung pada Lokasi 1 (Bogor-Cigombang) dan data pada Lokasi 12 (Boyolali-Bonyodono) memiliki standar deviasi yang relatif berbeda sehingga diputuskan untuk memisahkan data pada dua lokasi ini dari analisis. Setelah dilakukan pemisahan, dilakukan pengujian ulang dengan diperoleh nilai statistik uji (t-student) 1.58 dengan peluang nyata 0.071. Nilai ini menunjukkan bahwa ragam daya hasil untuk ke 16 lokasi memenuhi asumsi kehomogenan ragam. Selanjutnya pengujian normalitas untuk residual dengan menggunakan statistik uji Shapiro-Wilks sebesar 0.997 dengan nilai peluang nyata sebesar 0.374. Hasil ini menunjukkan data memenuhi asumsi kenormalan. Sehingga, untuk data dari 16 lokasi sudah memenuhi asumsi kehomogenan ragam dan kenormalan.

Dari hasil analisis ragam untuk daya hasil jagung dari 12 genotip pada 16 lokasi menunjukkan bahwa seluruh pengaruh utama (genotip dan lokasi) dan pengaruh interaksi genotip dengan lokasi berpengaruh nyata pada nilai peluang 0.0000. Hasil ini menunjukkan bahwa tingkat produksi jagung sangat dipengaruhi oleh faktor genotip dan lokasi. Jika dilihat dari sumbangan keragaman yang diberikan oleh masing-masing pengaruh terlihat pengaruh lokasi merupakan penyumbang keragaman produksi jagung terbesar, kemudian diikuti oleh faktor genotip kemudian interaksi genotip dengan lokasi. Dengan demikian tingkat produksi jagung akan sangat bergantung lokasi dimana jagung itu ditanam dan juga ditentukan oleh jenis genotip apa yang ditanam. Hasil ini juga terlihat dari respon lokasi yang sangat berfluktuatif yaitu berkisar 4.8588 ton/ha sampai 13.1254 ton/ha. Sedangkan respon genotip relatif seragam yaitu berkisar antara 8.887 ton/ha sampai dengan 10.126 ton/ha.

Tabel 4. Hasil Analisis Ragam Untuk Data Produksi Jagung

Model AMMI Kini dan Yang Akan Datang

Sumber Keragaman	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengan	F-Hitung	P
Genotip (G)	11	1.92	0.1742	11.1	0.00000
Lokasi (L)	15	66.29	4.4196	79.22	0.00000
Interaksi (G*L)	165	5.79	0.0351	2.24	0.00000
KUI 1	25	1.9	0.0761	4.85	0.00000
KUI 2	23	1.09	0.0475	3.03	0.00001
KUI 3	21	0.84	0.0399	2.54	0.00026
KUI 4	19	0.65	0.0343	2.19	0.00304
KUI 5	17	0.47	0.0279	1.78	0.02920
Simpangan	60	0.83	0.0138	0.88	0.72302
Galat	384	7.3075	0.01903	*	*
Total Terkoreksi	575	81.31	0.1414	*	*
Koef. Determinasi (R^2) = 0.9101 Koef. Keragaman (KK) = 12.49 Simpangan Baku galat (s) = 0.137949					

Secara deskriptif diketahui bahwa Lokasi 17 (Propinsi Jawa Timur, Kabupaten Jember Kecamatan Ambulu, Desa Pontang pada musim kemarau) yang memiliki tingkat produksi terbaik dengan tingkat produksi rata-rata mencapai 13,1254 ton/ha. Produksi terendah juga terdapat di Lokasi 18 (Propinsi Jawa Timur namun di kabupaten Malang, Kecamatan Tajinan desa Jambu Timur, pada musim kemarau) dengan rata-rata produksi jagung hanya sebesar 4.8588 ton/ha. Sedagkan genotip yang memberikan respon terbaik adalah BC 42521 dan yang terendah adalah genotip BIO 9900.

Tabel 4. menunjukkan bahwa interaksi antara jenis genotip dan lokasi tanam berpengaruh nyata terhadap produksi. Hasil tersebut berarti jenis genotip tertentu akan tumbuh baik pada lokasi tertentu tetapi tidak begitu halnya jika ditanam pada lokasi yang lain. Secara deskripsi dapat digambarkan bahwa untuk genotip BIO 9900 memberikan respon terbaik pada Lokasi -17 (Propinsi Jawa Timur, Kabupaten Jember Kecamatan Ambulu, Desa Pontang pada musim kemarau) namun kurang baik di Lokasi 18 (Jawa Timur, Kecamatan Tajinan, Desa Jambu Timur pada Musim Kemarau). Begitu juga untuk genotip BC 42521 juga memberikan respon yang tinggi pada Lokasi 17 (Propinsi Jawa Timur, Kabupaten Jember Kecamatan Ambulu, Desa Pontang pada musim kemarau) dan kurang baik pada Lokasi 18 (Jawa Timur, Kecamatan Tajinan, Desa Jambu Timur pada Musim Kemarau)

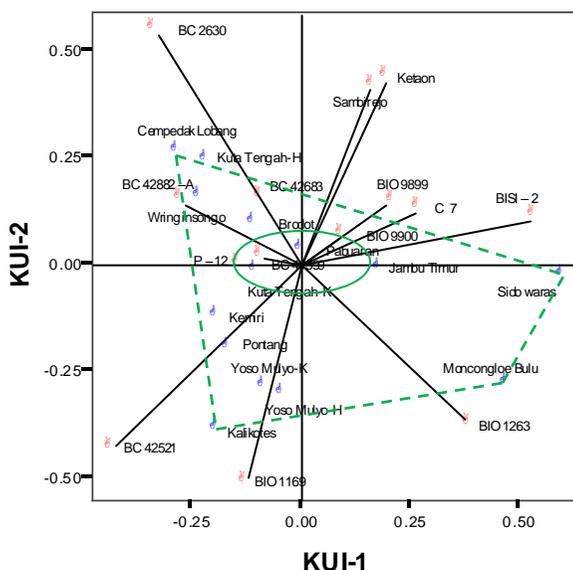
Jika dilihat dari hasil deskripsi di atas, khususnya untuk hasil produksi menurut lokasi, produksi rata-rata produksi jagung tertinggi dan terendah ada di propinsi Jawa Timur. Melihat fenomena ini tentunya dimungkinkan adanya faktor-faktor lingkungan yang berpengaruh terhadap produksi jagung semisal tinggi lokasi dan juga musim. Untuk kajian lebih lanjut analisis interaksi genotip dengan lokasi

Model AMMI Kini dan Yang Akan Datang

yang melibatkan pengaruh faktor lingkungan dan juga *genotypic* akan di bahwa pada topik selanjutnya yaitu *SEM-AMMI*.

Analisis AMMI

Penguraian bilinear terhadap matriks pengaruh interaksi dari data produksi jagung diperoleh 11 komponen utama interaksi (KUI) dengan nilai akar ciri masing-masing KUI adalah 0.634471, 0.364146, 0.279313, 0.217495, 0.158145, 0.104916, 0.060365, 0.048517, 0.031275, 0.021437, dan 0.009482. Kontribusi keragaman yang diterangkan oleh masing-masing komponen adalah 32.88%, 18.87%, 14.48%, 11.27%, 8.20%, 5.44%, 3.13%, 2.51%, 1.62%, 1.11%, dan 0.49%. Komponen utama interaksi yang nyata diperoleh dengan metode *Postdictive Success* menghasilkan lima komponen yang nyata dengan nilai F serta nilai peluangnya nyatanya masing-masing dapat dilihat pada Tabel 4. Hal ini berarti daya hasil dapat diterangkan dengan menggunakan model AMMI5, dimana pengaruh interaksi direduksi menjadi empat komponen. Dengan demikian model AMMI5 dapat menerangkan pengaruh interaksi sebesar 85,7%. Untuk keperluan analisis stabilitas dan daptabilitas, digunakan Biplot AMMI2 beserta selang kepercayaan elips. Biplot AMMI2 merupakan plot antara KUI1 dengan KUI2.



Gambar 2. Biplot AMMI2 Daya Hasil (51.75%)

Perhitungan persamaan elips untuk daya hasil menghasilkan jari-jari panjang 0.153 dan jari-jari pendek 0.092. Terlihat genotip BIO 9900, BC 41399 dan P - 12 terletak di dalam ellips dengan genotip BC 41399 yang paling dekat dengan titik pusat (0.0). Hasil ini menunjukkan bahwa ketiga genotip tersebut stabil dengan

genotip paling stabil adalah BC 41399. Biplot AMMI2 menunjukkan genotip yang berinteraksi khas dengan lokasi tertentu. Makin dekat jarak lokasi dengan genotip, atau semakin kecil sudut diantara keduanya maka makin kuat interaksinya.

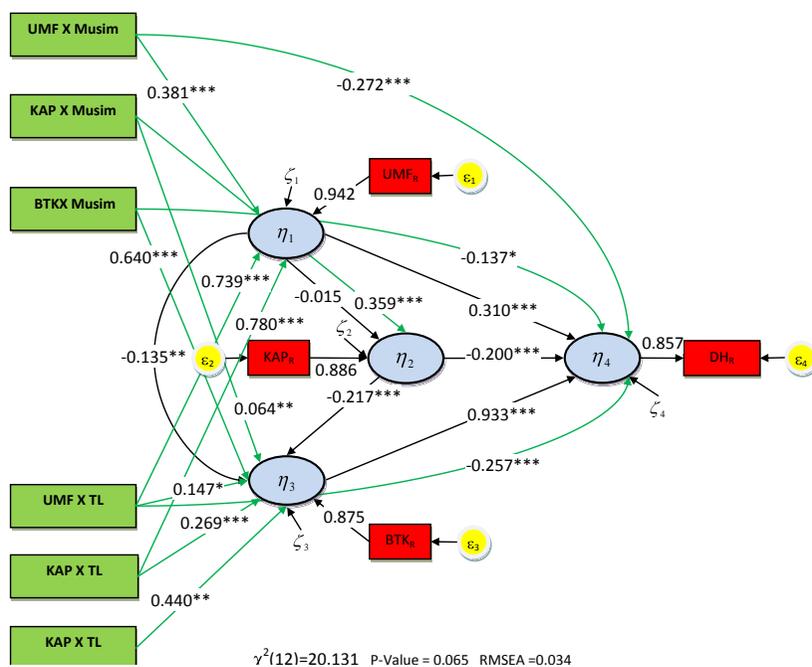
Gambar 2 menunjukkan bahwa genoti-genotip yang spesifik pada lokasi tertentu yaitu genotip BIO1169 dan BC 42521 pada lokasi L13 (Kalikotes-Musim Kemarau). Genotip BIO 1263 pada lokasi 4 (Moncongloe Bulu-Musim Hujan). Selanjutnya Genotip BISI-2 pada lokasi 6 (Sidowaras – Musim Hujan). Genotip BC 2630 cocok pada lokasi 14 (Kuta Tengah – Musim Hujan) dan juga pada Lokasi L10 (Cempedak Lobang – Musim Hujan). Genotip BC 42882 cocok pada lokasi L8 (Wringin Solo – Musim Hujan).

Structural equation modeling- AMMI

Struktural Equation Modeling merupakan satu teknik analisis statistik yang digunakan untuk mengkaji hubungan sebab akibat antara variabel penelitian yang memungkinkan memasukkan kekeliruan pengukuran dalam pemodelan suatu system persamaan. Melalui teknik ini juga dapat dilakukan analisis suatu sistem persamaan simultan dan memungkinkan melakukan pengujian koefisien arah secara parsial. Analisis *Struktural Equation Modeling* dicoba untuk dikombinasikan dengan pemodelan *AMMI* untuk membantu memberikan penjelasan faktor-faktor yang mempengaruhi signifikansi efek interaksi genotip \times lokasi dimana faktor-faktor tersebut adalah komponen-komponen dari daya hasil dan juga faktor *genotypic* dan lingkungan yang selama ini dalam analisis interaksi genotip \times lokasi sering terabaikan, dimana faktor-faktor tersebut dinilai memiliki peran penting dalam menjelaskan interaksi genotip \times lingkungan khususnya dalam menemukan varietas unggulan.

Seperti yang telah dijelaskan pada bab sebelumnya bahwa dalam analisis *SEM* dilibatkan variabel eksogen dan variabel endogen. Variabel eksogen dalam kasus ini adalah kovariat *genotypic* \times lokasi, sedangkan variabel endogennya adalah *GEI* untuk komponen daya hasil dan *GEI* untuk daya hasil. Komponen daya hasil yang dilibatkan adalah umur masak fisiologis, kadar air panend dan berat tongkol.

Model AMMI Kini dan Yang Akan Datang



Gambar 3. Diagram Lintas SEM-AMMI

Skor untuk *GEI* yang merupakan skor faktor laten endogen dalam *SEM* diperoleh dari dekomposisi singular matriks interaksi yang merupakan nilai yang diperoleh dengan menghilangkan efek utama dalam analisis ragam. Melalui dekomposisi ini diperoleh skor *full* interaksi yang telah dipisahkan dari komponen *noise*. Skor *GEI* untuk komponen daya hasil dan daya hasil diperoleh dari lima komponen utama dari hasil pengujian *Postdictive Success*. Dari lima komponen utama yang terpilih dapat menjelaskan 94,2%, 88,6%, 87,5% dan 85,7% pengaruh interaksi untuk masing-masing *GEI* pada umur masak fisiologis (UMF), kadar air panen (KAP), berat tongkol (BTK) dan daya hasil (DY). Variabel eksogennya merupakan faktor *genotypic* \times lingkungan diperoleh dari perkalian silang nilai komponen residual dengan nilai faktor lingkungan kemudian disesuaikan terhadap efek utama seperti pada penyesuaian untuk faktor genotipe.

Melalui pendugaan dengan metode ADF menggunakan software AMOS 7 diperoleh model fit (*closed fit*) dengan data atau *covariance matriks observation fit* dengan *covariance matriks* prediksi yang ditunjukkan dengan nilai peluangnya nyata ChiSquare ($\chi^2(12)=20.131$, p-value=0.065, RMSEA = 0.034) lebih kecil dari 0.05, dan juga nilai GFI, AGFI, NFI yang lebih besar dari 0.90. Pengujian koefisien jalur secara parsial menunjukkan bahwa hanya koefisien jalur dari *GEI* untuk umur masak fisiologis terhadap *GEI* untuk umur kadar air panen yang tidak signifikan.

Tabel 5. Nilai Kecocokan Model

Goodness of Fit Statistics			
Absolute fit measure	Nilai		Keterangan
Degrees of Freedom	12		Model Fit
Minimum Fit Function Chi-Square	20.131 (P = 0.065)	P<0.05	Model Fit
Goodness of Fit Index (GFI)	0.988	> 0.90	Model Fit
Root Mean Square Error of Approximation (RMSEA)	0.034	<0.05	Model Fit

Pemilihan variabel eksogen yaitu kovariate *genotypic* × lokasi yang dimasukkan ke dalam model dilakukan dengan prosedur *stepwise* yaitu memasukkan satu-persatu variabel eksogen ke dalam model kemudian dilakukan evaluasi model. Jika dengan memasukkan variabel eksogen tertentu variabel tersebut signifikan, maka variabel tersebut dipertahankan dalam model jika tidak variabel tersebut dikeluarkan dari model. Hasil analisis *SEM* diketahui bahwa signifikansi efek interaksi genotip x lokasi daya hasil dipengaruhi oleh *GEI* komponen daya hasil yaitu Umur Masak Fisiologis, Kadar Air Panen, dan Berat Tongkol dengan berat tongkol memberikan efek langsung terbesar dengan *standardized* efek masing-masing adalah 0.310, -0.200, 0.933. *GEI* untuk komponen daya hasil Umur Masak fisiologis memiliki efek langsung positif terhadap *GEI* daya hasil. Ini artinya bahwa untuk Umur Masak fisiologis di atas rata-rata maka akan memiliki daya hasil yang relative lebih banyak. Begitu juga *GEI* untuk berat tongkol. Namun, untuk *GEI* untuk komponen daya hasil Kadar Air Panen memiliki efek langsung negatif terhadap *GEI* untuk daya hasil. Hasil ini memberikan informasi bahwa jika kadar air panen di atas rata-rata maka daya hasil atau produksi jagung relative lebih sedikit. *GEI* untuk umur masak fisiologis yang memberikan efek tidak langsung melalui kadar air panen, dan berat tongkol sebesar -0.090. Tanda negatif ini terjadi karena melalui kadar air panen yang memiliki efek negative pada daya hasil. Selanjutnya, kadar air panen juga memberikan efek tidak langsung terhadap *GEI* untuk daya hasil melalui berat tongkol dengan besar efek tidak langsungnya adalah -0.209. Total efek dari ketiga *GEI* untuk komponen daya secara berurutan adalah 0.241 dari *GEI* untuk Uurm Masak Fisiologis, -0.413 dari *GEI* Kadar Air Panen, dan 0.921 dari *GEI* untuk berat tongkol. Dari model *SEM* ini juga dapat diketahui keragaman dari *GEI* untuk Umur Masak Fisiologis, *GEI* untuk Kadar Air Panen, dan *GEI* untuk berat tongkol secara berurutan adalah 0.874, 0.812, dan 0.767. Besarnya nilai ini menunjukkan bahwa model yang dianalisis dapat menjelaskan keterkaitan antara *GEI* untuk komponen daya hasil dan

menjelaskan pengaruh faktor-faktor *genotypic* × lokasi terhadap *GEI* komponen daya hasil.

Kovariat *genotypic* × lokasi yang berpengaruh signifikan terhadap daya hasil adalah Umur Masak Fisiologi × Tinggi Tempat, Umur Masak Fisiologis × Musim, dan Berat Tongkol Musim. Hasil ini memberikan gambaran bahwa jagung dengan usia masak fisiologis yang relatif lama ditanam pada lingkungan dengan lokasi relatif tinggi maka akan berakibat pada daya hasil yang kurang baik. Begitu juga jika di tanam pada musim dimana hujannya relatif tinggi. Sedangkan jika untuk genotip jagung yang memiliki berat tongkol relatif berat jika di tanam pada lokasi dengan tingkat curah hujan yang tinggi maka daya hasil jagung akan relatif kurang baik. Secara keseluruhan, model *SEM* ini dapat menjelaskan keragaman dari *GEI* daya hasil sebesar 0.711. Nilai ini berarti bahwa faktor-faktor genotip × lingkungan khususnya umur masak fisiologis × tinggi lokasi, umur masak fisiologis × musim dan berat tongkol × musim, beserta *GEI* komponen daya memiliki peran penting dalam mempengaruhi daya hasil.

KESIMPULAN

Melalui *Struktural Equation Modeling* dapat dilakukan analisis *GEI* untuk daya hasil dengan memasukkan komponen daya hasil dan kovariat *genotypic* × lokasi yang berguna untuk mengidentifikasi faktor-faktor yang menyebabkan nyataanya efek interaksi genotip × lokasi dalam analisis Ragam *AMMI*.

Hasil analisis untuk data produksi jagung dengan 12 genotip yang ditanam pada 16 lokasi diketahui bahwa signifikansi efek interaksi genotip × lokasi dipengaruhi oleh *GEI* komponen daya hasil yaitu Umur Masak Fisiologis, Kadar Air Panen, dan Berat Tongkol dengan berat tongkol memberikan efek langsung terbesar. *GEI* untuk komponen daya hasil Umur Masak fisiologis memiliki efek langsung positif terhadap *GEI* daya hasil. Ini artinya bahwa untuk Umur Masak fisiologis di atas rata-rata maka akan memiliki daya hasil yang relative lebih banyak. Begitu juga *GEI* untuk berat tongkol. Namun, untuk *GEI* untuk komponen daya hasil Kadar Air Panen memiliki efek langsung negative terhadap *GEI* untuk daya hasil. Hasil ini memberikan informasi bahwa jika kadar air panen di atas rata-rata maka daya hasil atau produksi jagung relative lebih sedikit.

Kovariat *genotypic* × lokasi yang berpengaruh signifikan terhadap daya hasil adalah Umur Masak Fisiologi × Tinggi Tempat, Umur Masak Fisiologi × Musim, dan Berat Tongkol Musim. Hasil ini memberikan gambaran bahwa jagung dengan usia masak fisiologis yang relatif lama ditanam pada lingkungan dengan lokasi

relatif tinggi maka akan berakibat pada daya hasil yang kurang baik. Begitu juga jika di tanam pada musim dimana hujannya relatif tinggi. Sedangkan jika untuk genotip jagung yang memiliki berat tongkol relatif berat jika di tanam pada lokasi dengan tingkat curah hujan yang tinggi maka daya hasil jagung akan relatif kurang baik.

UCAPAN TERIMA KASIH

Tulisan ini bagian dari Hibah Penelitian Tim Pascasarjana yang didanai oleh Direktorat Jenderal Pendidikan Tinggi Nomor : 226/13.11/PL/2008 Tanggal : 02 April 2008.

DAFTAR PUSTAKA

- Aastveit H., and H. Martens. 1986. ANOVA Interactions Interpreted by Partial Least Squares Regression. *Biometrics* 42 : 829-844
- Alberts Martin J. A. 2004. *A Comparison of Statistical Method To Describe Genotype \times Environment Interaction and Yield Stability In Multilocation Maize Trials*. Thesis. University of The Free State Bloemfontein
- Arbuckle. J.L. 2006. Amos™ 7.0 User's Guide, Amos Development Corporation, United States of America.
- Bollen, K.A. 1989. *Structural Equation With Latent Variables*. New York: John Wiley and Sons.
- Cambell, B. T. 2002. *Quantitative Trait Loci and Environmental Interactions Associated With Agronomic Performance of Wheat*. Ph.D. Dissertation. Nebraska: University of Nebraska-Lincoln.
- Dhungana, P. 2004. *Structural Equation Modeling of Genotype X Environment Interaction*. Ph. D. Dissertation. University of Nebraska-Lincoln, Lincoln, Nebraska.
- Ghozali, Imam. 2004. *Model Persamaan Structural Konsep dan Aplikasi dengan Program AMOS Versi 5.0*. Semarang: Badan Penerbit Universitas Diponegoro,
- Gauch, H.G. JR. 1982. Noise reduction by eigenvector ordination. *Ecology* 63:1643-1649
- Gauch, H.G. JR. 1982. Model selection and validation for yield trials with interaction. *Biometrics* 44: 705-715
- Joreskog, K. G. & Sorbom, D. 1989. LISREL 7 : User's reference guide. Scientific Software Inc., Mooresville, IN.
- Johnson, R.A. & D.W. Wichern. 1996. *Applied Multivariate Statistical Analysis 3rd ed*. New Jersey: Prentice-Hall.

Model AMMI Kini dan Yang Akan Datang

Mattjik, A.A. & Sumertajaya, I.M. 2000. *Perancangan Percobaan dengan Aplikasi SAS dan Minitab* jilid 1. Bogor: IPB Press.

Tai, G.C.C. 1979. Analysis of Genotype-Environment Interactions of Potato Yield. *Crop Sci.* 19:434-438

_____2002. *Manual SEM-AMOS.*

7.2 ANALISIS INTERAKSI GENOTIPE-LINGKUNGAN MENGGUNAKAN PARTIAL LEAST SQUARE PATH MODELING (I Made Sumertajaya, Jurnal, BIAStatistika Vol. 4, No. 1, hal. 9-18)

ANALISIS INTERAKSI GENOTIPE-LINGKUNGAN MENGGUNAKAN PARTIAL LEAST SQUARE PATH MODELING

I Made Sumertajaya

Departemen Statistika Institut Pertanian Bogor

ABSTRAK

Percobaan multilokasi telah banyak digunakan oleh para pemulia tanaman untuk mengkaji kemampuan relatif genotipe-genotipe pada berbagai Lokasi tanam dengan tujuan menemukan genotipe-genotipe unggulan. Nyatanya pengaruh interaksi genotipe \times lokasi (IGL) pada percobaan multilokasi menyulitkan dalam proses seleksi genotipe unggulan. Sehingga memahami faktor lokasi dan genotipik yang berpengaruh terhadap nyatanya IGL akan sangat membantu pada setiap tahapan pemuliaan tanaman. Tujuan dari penelitian ini adalah mengkaji struktur interaksi dari karakteristik agronomi tanaman jagung menggunakan kombinasi metode Additive Main Effect Multiplicative Interaction (AMMI) dengan model persamaan struktural. Penggabungan model persamaan struktural dengan AMMI memungkinkan peneliti menjelaskan interaksi genotipe \times lokasi dengan memasukkan informasi rangkaian proses biologis yang terkait dalam pertumbuhan dan perkembangan tanaman serta memasukkan informasi kombinasi kovariat genotipik dan lokasi. Dalam penelitian ini, peneliti menggunakan data hasil pemuliaan jagung hibrida dengan 9 genotipe harapan dan 3 genotipe komersial yang dicobakan pada 16 lokasi. Karakteristik agronomi yang diamati sesuai dengan kajian literatur adalah usia masak fisiologis (UMF), kadar air panen (KAP), berat tongkol panen (BTK), dan hasil (HSL). Hasil eksplorasi data menunjukkan data tidak menyebar normal multivariat sehingga dalam pemodelan persamaan struktural menggunakan pendekatan partial least square (PLS). Hasil analisis Biplot-AMMI memberikan kesimpulan bahwa genotipe-genotipe yang relatif stabil adalah genotipe BC 41399 (F), BIO 9899 (I) dan BC 42683 (E) dan faktor yang paling berpengaruh terhadap hasil jagung hibrida adalah berta tongkol panen.

Kata Kunci : Model AMMI, Model Persamaan Struktural, PLS, Biplot-AMMI

1. PENDAHULUAN

Seleksi genotipe unggulan seringkali sulit dilakukan karena nyatanya efek interaksi genotipe \times lokasi (IGL) dalam percobaan multilokasi. Dibutuhkan pemahaman yang lebih dalam tentang IGL agar dapat membantu proses seleksi. Kajian tentang IGL telah banyak dilakukan diantaranya menggunakan metode *Additive Main Effect Multiplicative Interaction (AMMI)* (Gauch, 1988). Namun sedikit kajian yang menjelaskan bagaimana pengaruh komponen hasil dan kombinasi kovariat genotipik \times lokasi terhadap nyatanya efek IGL hasil. Kajian ini diperlukan untuk mengidentifikasi pada kondisi lokasi dan karakteristik seperti apa genotipe akan memberikan hasil yang lebih baik, karena hasil adalah akumulasi respon dari karakteristik genotipe terhadap kondisi lokasi selama proses pertumbuhan dan perkembangannya (Dhungana, 2004). Dengan kata lain, kajian ini memberikan informasi awal kepada pemulia tanaman untuk lebih fokus pada karakteristik genotipe dan faktor lokasi yang paling berperan dalam peningkatan hasil.

Penggabungan metode *Partial least square path modeling* (PLSPM) dengan AMMI diperkenalkan untuk memodelkan IGL hasil dengan memasukkan variabel komponen hasil dan kombinasi kovariat genotipik \times lokasi sehingga dapat memberikan penjelasan secara komprehensif mengenai nyatanya efek IGL hasil. Melalui AMMI diperoleh variabel laten IGL komponen-komponen hasil dan hasil yang telah terbebas dari cecaran galat (*noise*). Sedangkan melalui pemodelan persamaan struktural yaitu melalui *partial least square* akan dapat dibuat sebuah sistem persamaan yang menjelaskan keterkaitan antara variabel laten IGL dengan memasukkan informasi kovariat lokasi \times genotipik. Pendekatan *Partial Least Square-AMMI* dalam pemodelan IGL hasil akan memberikan informasi yang akurat karena variabel laten yang digunakan telah dimurnikan dari cecaran galat (*noise*) dan memberikan informasi kecocokan model (*goodness of fit*) sebagai indikator kemampuan model dalam menjelaskan keragaman data. Kajian ini pernah dilakukan oleh Dhungana (2004) menggunakan data padi dengan asumsi data menyebar mengikuti sebaran normal ganda. Dalam penelitian ini PLS-AMMI akan diterapkan pada data Jagung dengan kondisi data tidak menyebar normal ganda.

2. TUJUAN

1. Mengkaji struktur pengaruh interaksi dengan pendekatan model AMMI (Additive Main Effects and Multiplicative Interaction) untuk komponen hasil dan hasil.
2. Menjelaskan interaksi genotipe \times lokasi menggunakan *partial least square path modeling*

3. METODOLOGI

3.1 Data Penelitian

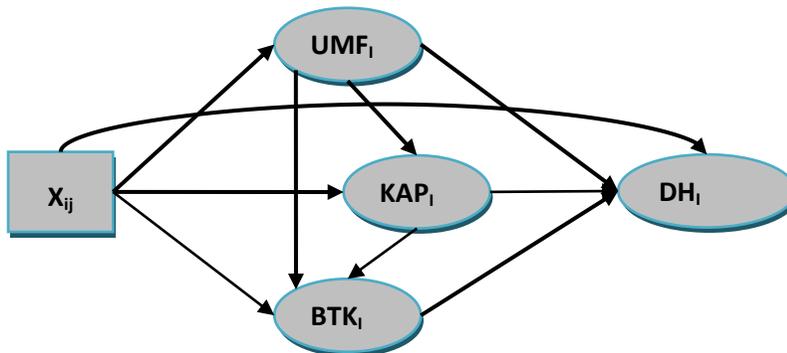
Data yang digunakan dalam penelitian ini adalah data hasil percobaan multilokasi Jagung Hibrida yang dilakukan dari tanggal 23 Juli 2006 sampai 10 April 2007 yaitu pada musim hujan dan kemarau. Percobaan ini menggunakan 9 genotipe Jagung Hibrida Harapan dan 3 genotipe Jagung Hibrida Komersial. Penelitian ini mengambil 16 lokasi tanam yang tersebar di 6 Propinsi di Indonesia. Percobaan multilokasi dilakukan dengan rancangan acak kelompok (RAK) dengan kelompok tersarang pada lokasi.

3.2 Metode Penelitian

Dalam penelitian ini tahapan analisisnya adalah sebagai berikut :

1. Menetapkan model konseptual dari IGL Hasil.

Model koseptual ditetapkan berdasarkan kajian literatur dan eksplorasi data dengan model yang akan diuji adalah :



Gambar 1. Model Penelitian

Keterangan : Skor Interaksi Usia Masak Fisiologis (UMF_i), Skor Interaksi Kadar Air Saat Panen (KAP_i), Skor Interaksi Berat tongkol panen (BTK_i), Skor Interaksi Hasil (DH_i), dan Kovariat genotipik \times lokasi (X_{ij})

2. Analisis struktur interaksi komponen hasil dan struktur interaksi hasil menggunakan metode AMMI.

Pemodelan Analisis AMMI

Langkah awal untuk memulai analisis AMMI adalah melakukan analisis ragam gabungan untuk mengetahui apakah IGL nyata. Selanjutnya struktur IGL dijelaskan menggunakan analisis komponen utama. Bentuk multiplikatif

Model AMMI Kini dan Yang Akan Datang

diperoleh dari penguraian interaksi genotipe dengan lokasi menjadi komponen utama interaksi (KUI).

Model analisis ragam untuk AMMI menggunakan rancangan acak kelompok dengan kelompok tersarang pada lokasi sebagai berikut :

$$\begin{aligned}y_{ger} &= \mu + \alpha_g + \beta_e + \theta_{r|e} + \gamma_{ge} + \varepsilon_{ger} \\ &= \mu + \alpha_g + \beta_e + \theta_{r|e} + \sum_{j=1}^n \lambda_j u_{gj} v_{ej} + \delta_{ge} + \varepsilon_{ger} \\ y_{ger} &= \mu + \alpha_g + \beta_e + \lambda_1 u_{g1} v_{e1} + \lambda_2 u_{g2} v_{e2} + \dots + \lambda_n u_{gn} v_{en} + \delta_{ge} + \varepsilon_{ger} \quad (1)\end{aligned}$$

dengan $g = 1, 2, \dots, a$; $e = 1, 2, \dots, b$; $n = 1, 2, \dots, k$ dan $k =$ banyaknya KUI yang nyata pada taraf 5% .

selanjutnya :

- μ : nilai rata-rata umum
- α_g : pengaruh utama genotipe ke- g
- β_e : pengaruh utama lokasi ke- e
- γ_{ge} : pengaruh interaksi genotipe ke- g dengan lokasi ke- e
- ε_{ger} : pengaruh acak pada genotipe ke- g , lokasi ke- e dan ulangan ke- r
- λ_n : nilai singular untuk komponen bilinear ke- n
- u_{gn} : pengaruh genotipe ke- g melalui komponen bilinear ke- n
- v_{en} : pengaruh lokasi ke- e melalui komponen bilinear ke- n ,
- δ_{ge} : simpangan dari pemodelan bilinear

Untuk identifikasi genotipe stabil dan spesifik lokasi digunakan biplot dan untuk menentukan peringkat genotipe stabil dilakukan dengan formulasi *Indeks Stabilitas AMMI* yang penulis kembangkan dari konsep phytagoras dalam biplot.

$$ISA = \sqrt{\left(\left[\frac{\lambda_1^{1/2}}{\lambda_2^{1/2}} (\text{SkorKU} - 1) \right]^2 + \left[\text{SkorKU} - 2 \right]^2 \right)} \quad (2)$$

Penguraian Nilai Singular (SVD=Singular Value Decomposition)

Penguraian nilai singular matriks interaksi Γ digunakan untuk menduga pengaruh interaksi genotipe x lokasi. Penguraian dilakukan dengan memodelkan matriks tersebut sebagai perkalian matriks :

$$\Gamma = \underset{nxp}{\mathbf{U}} \underset{nxk}{\mathbf{\Omega}} \underset{kxk}{\mathbf{V}}', \quad (3)$$

dengan elemen matriks Γ dihitung menggunakan formulasi :

$$\gamma_{ge} = \bar{y}_{ge.} - \bar{y}_{g..} - \bar{y}_{.e.} + \bar{y}_{...} \quad (4)$$

3. Mendapatkan variabel latent IGL komponen hasil dan variabel laten IGL untuk hasil serta mengoreksi kovariat genotipe x lokasi terhadap pengaruh utama.

Misalkan $\mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3,$ dan \mathbf{Y}_4 masing-masing adalah matriks interaksi $DH_i, UMF_i, KAP_i,$ dan BTK_i dengan ordo masing-masing $n \times p$ dengan n adalah banyaknya genotipe dan p adalah banyaknya lokasi. Setiap matriks interaksi genotipe x lokasi dapat didefinisikan menggunakan *singular value decomposition* (SVD) sebagai berikut :

$$\mathbf{Y}_i = \underset{(nxp)}{\mathbf{U}}_i \underset{(nxk)}{\mathbf{\Omega}}_i \underset{(kxp)}{\mathbf{V}}_i' + \mathbf{e}_i, \quad (5)$$

Diasumsikan bahwa $\mathbf{U}_i \mathbf{\Omega}_i \mathbf{V}_i'$ adalah nilai IGL sebenarnya dari variabel ke- i dengan k komponen pertama ditentukan berdasarkan pada metode keberhasilan total (*postdictive success*). Matriks \mathbf{Y}_i dalam persamaan (5) dikonversi ke vektor kolom dengan menggunakan operator **vec** dan produk **konecker** (Harville, 1997) :

$$\mathbf{Vec}(\mathbf{Y}_i) = (\mathbf{V}_i \otimes \mathbf{U}_i) \mathbf{vec}(\mathbf{\Omega}_i) + \mathbf{vec}(\mathbf{e}_i) \quad (6)$$

$(np \times 1)$ $(np \times kk)$ $(kk \times 1)$ $(np \times 1)$

$$\text{dan } \eta_i = (\mathbf{V}_i' \otimes \mathbf{U}_i) \mathbf{vec}(\mathbf{\Omega}_i), \quad (7)$$

4. Tahap selanjutnya adalah pemodelan IGL dengan *Partial Least Square*

Partial Least Square (PLS)

Metode Struktural Equation Modeling berbasis komponen sering disebut dengan metode *Partial Least Square* (PLS). Metode *Partial Least Square* dikembangkan pertama kali oleh Wold dalam Chin (2000) sebagai metode umum untuk pendugaan model peubah laten yang diukur tidak langsung oleh peubah penjelas. Metode ini dikembangkan sebagai solusi dari pelanggaran-pelanggaran asumsi yang dibutuhkan pada saat pemodelan persamaan structural dengan menggunakan LISREL. Metode PLS didefinisikan dari dua persamaan linier yang disebut model structural dan model pengukuran.

Model Structural

$$\eta_j = \beta_{j0} + \sum_{i < j} (\beta_{ji} \eta_i) + \zeta_j, \text{ untuk } j = 1, \dots, J \quad (5)$$

dengan :

J = banyaknya peubah latent

η_j = peubah latent tak bebas ke-j

η_i = peubah latent tak bebas ke-i untuk $i \neq j$

β_{ij} = koefisien jalur peubah laten ke-j dan ke-i

β_{i0} = intrsep

ζ_j = sisaan model structural ke-j

i = banyaknya jalur dari peubah laten bebas ke peubah latent tak bebas.

Model Pengukuran

$$y_{kj} = \pi_{kj0} + \pi_{kj}\eta_j + \varepsilon_{kj}, \quad \text{untuk } j=1,\dots,J, \quad k=1,\dots,K \quad (6)$$

dengan :

J = banyaknya peubah latent

η_j = peubah latent tak bebas ke-j

y_{kj} = peubah penjelas ke-k dan peubah laten ke-j

π_{kj} = koefisien antara peubah penjelas ke-k dan peubah laten ke-j

π_{kj0} = intrsep

ε = sisaan model pengukuran peubah penjelas ke-k dan peubah latent ke-j

k = lintasan dari η_j ke y_{kj}

K = banyaknya peubah penjelas pada peubah laten ke-j

Metode Estimasi

Metode pendugaan parameter (estimasi) di dalam PLS adalah metode kuadrat terkecil (*least square methods*). Proses perhitungan dilakukan dengan cara iterasi, dimana iterasi akan berhenti jika telah tercapai kondisi konvergen.

Pendugaan parameter di dalam PLS meliputi 3 hal, yaitu (Chin,2000):

- 1) *Weight estimate* yang digunakan untuk menciptakan skor variabel laten
- 2) Estimasi jalur (*path estimate*) yang menghubungkan antar variabel laten dan estimasi *loading* antara variabel laten dengan indikatornya.
- 3) *Means* dan lokasi parameter (nilai konstanta regresi, intersep) untuk indikator dan variabel laten.

Goodness of Fit Model

Goodness of Fit Model diukur menggunakan R-square variabel laten dependen dengan interpretasi yang sama dengan regresi; *Q-Square predictive relevance* untuk model struktural, mengukur seberapa baik nilai observasi dihasilkan oleh

model dan juga estimasi parameternya. Nilai Q-square > 0 menunjukkan model memiliki *predictive relevance*; sebaliknya jika nilai Q-Square ≤ 0 menunjukkan model kurang memiliki *predictive relevance*. Perhitungan Q-Square dilakukan dengan rumus:

$$Q_2 = 1 - (1 - R_1^2) (1 - R_2^2) \dots (1 - R_p^2) \tag{7}$$

dimana $R_1^2, R_2^2 \dots R_p^2$ adalah R-square variabel endogen dalam model persamaan. Besaran Q^2 memiliki nilai dengan rentang $0 < Q^2 < 1$, dimana semakin mendekati 1 berarti model semakin baik. Besaran Q^2 ini setara dengan koefisien determinasi total pada analisis jalur (*path analysis*).

Uji Partial

Dalam Partial Least Square tidak disajikan uji partial sehingga pendekatan pengujian secara partial dapat dilakukan dengan metode Bootstrap ataupun Jackknife.

4. HASIL DAN PEMBAHASAN

Analisis Deskriptif Data

Tabel 2. Rataan Genotipe Untuk Komponen Daya Hasil dan Daya Hasil

Genotipe	UMF	Rank (UMF)	KAP	Rank (KAP)	BTK	Rank (BTK)	HASIL	Rank (Hasil)
BIO 9900	99.00	3	27.49	3	10.76	2	8.89	2
BIO 1263	99.58	4	27.58	4	11.02	4	9.02	4
BIO 1169	100.10	5	27.22	1	11.13	5	9.18	8
BC 42521	101.56	11	27.74	6	12.53	12	10.13	12
BC 42683	100.90	9	27.92	9	11.69	11	9.51	11
BC 41399	100.77	8	27.89	8	11.58	10	9.50	10
BC 2630	98.58	1	28.51	12	11.15	6	8.99	3
BC 42882 – A	100.65	7	28.31	11	11.25	8	9.16	7
BIO 9899	100.92	10	27.32	2	11.31	9	9.31	9
BISI – 2	102.54	12	27.93	10	10.43	1	8.69	1
P – 12	99.00	3	27.64	5	11.20	7	9.08	6
C -7	100.23	6	27.82	7	11.01	3	9.05	5

Di bawah ini disajikan data rataan genotipe untuk komponen daya hasil dan daya hasil. Dari Tabel 2 dapat diamati bahwa genotipe dengan daya hasil relatif tinggi adalah genotipe BC 42521, BC 42683, dan BC 41399. Namun perlu diingat bahwa karakteristik genotipe unggulan tidak hanya didasarkan pada daya hasil yang relatif tinggi namun memiliki tingkat adaptasi yang tinggi pada berbagai kondisi lingkungan. Selanjutnya akan dilakukan analisis ragam gabungan untuk

Model AMMI Kini dan Yang Akan Datang

identifikasi apakah IGL nyata untuk setiap komponen daya hasil dan daya hasil. Untuk memenuhi asumsi kenormalan galat dan kehomogenan ragam data yang digunakan adalah data hasil transformasi akar kuadrat.

Tabel 3. Analisis Ragam

Komponen Daya Hasil dan Daya Hasil	Jumlah Kuadrat					Persen (%)
	Genotipe	Lingkungan	Interaksi	Ulangan Lingkungan	Total	
Usia Masak Fisiologis (UMF)	1.746**	53.217**	5.166**	0.147**	60.834	8.492
Kadar Air Panen (KAP)	0.691**	32.468**	6.005**	1.194**	46.449	12.927
Berat Tongkol (BTK)	3.155**	162.358**	7.374**	2.868**	182.184	4.047
Daya Hasil (DH)	1.916**	66.295**	5.789**	1.785**	81.306	7.120

Keterangan : **) Nyata pada $\alpha=0.05$

Tabel 3 di atas menunjukkan bahwa terjadi interaksi yang nyata antara genotipe x lingkungan baik untuk komponen-komponen daya hasil dan daya hasil sendiri dengan persentase keragaman yang dijelaskan oleh komponen interaksi berkisar dari 4%-13%. Nyatanya efek interaksi ini menyatakan bahwa genotipe-genotipe memberikan respon yang berbeda untuk lingkungan yang berbeda. Perbedaan respon genotipe-genotipe ini tidak terlepas dari tingginya keragaman lingkungan makro untuk lokasi percobaan. Hasil ini memberikan gambaran bahwa faktor lingkungan memberikan pengaruh terhadap nyata IGL komponen daya hasil dan IGL daya hasil. Nyatanya efek interaksi ini menyulitkan dalam identifikasi genotipe-genotipe stabil sehingga diperlukan kajian lebih lanjut untuk identifikasi genotipe stabil dan spesifik lingkungan dengan AMMI dan Biplot.

Tabel 4. Proporsi Keragaman Interaksi

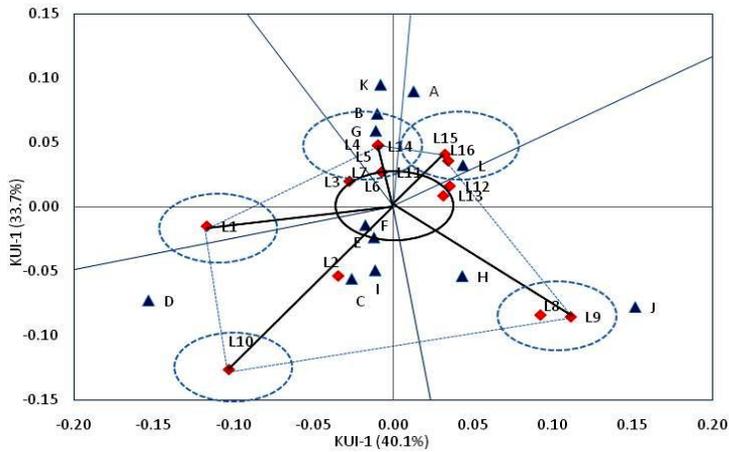
Komponen AMMI	DF	Usia Masak		Kadar Air Panen		Berat Tongkol		Daya Hasil	
		Prop	Kum	Proporsi	Kumulatif	Proporsi	Kumulatif	Proporsi	Kumulatif
KUI 1	25	0.402	0.402	0.354	0.354	0.366	0.366	0.329	0.329
KUI 2	23	0.337	0.739	0.177	0.531	0.201	0.567	0.189	0.518
KUI 3	21	0.086	0.825	0.138	0.669	0.163	0.73	0.145	0.662
KUI 4	19	0.075	0.9	0.116	0.785	0.079	0.809	0.113	0.775
KUI 5	17	0.043	0.942	0.101	0.886	0.066	0.875	0.082	0.857

Kajian struktur interaksi genotipe x lingkungan untuk komponen daya hasil dan daya hasil menunjukkan bahwa model AMMI nyata sampai pada KUI 5 dengan proporsi keragaman dari komponen interaksi yang dapat dijelaskan berkisar

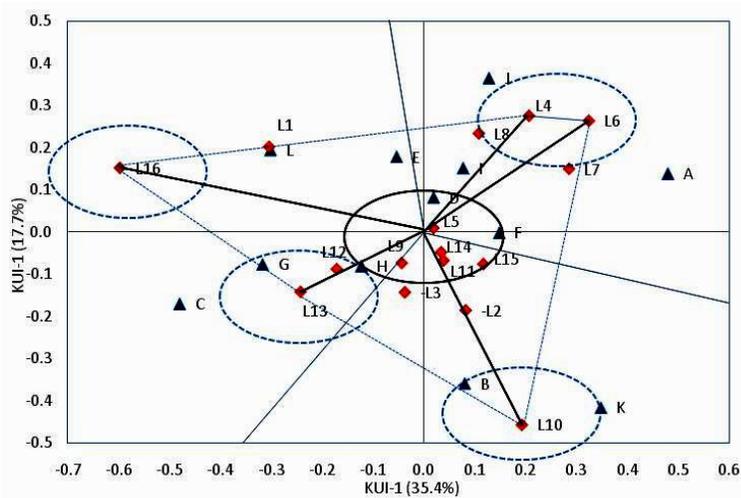
antara 85.7%-94.2% sehingga model AMMI untuk masing-masing komponen daya hasil dan daya hasil adalah :

$$\hat{y}_{ger} = \hat{\mu} + \hat{\alpha}_g + \hat{\beta}_e + \hat{\theta}_{rle} + \sum_{j=1}^5 \lambda_j u_{gj} v_{ej} + e_{ger} \quad (13)$$

Model ini dapat dijadikan model penduga nilai respon yang akurat karena telah mengeluarkan cemaran galat dari komponen interaksi melalui penguraian singular.



Gambar 2. Biplot UMF (73.9%)



Gambar 3. Biplot KAP (53.1%)

Model AMMI Kini dan Yang Akan Datang

Identifikasi genotipe stabil ini kurang baik jika hanya mempertimbangkan daya hasil semata. Namun sebelum memutuskan ketiga genotipe tersebut stabil harus ditunjukkan bahwa komponen-komponen daya hasil tersebut secara nyata merupakan indikator stabilitas daya hasil dengan PLSPM-AMMI.

Tabel 5. Indeks Stabilitas AMMI

Kode	Genotipe	UMF		KAP		BTK		DH	
		ASV	Rank	ASV	Rank	ASV	Rank	ASV	Rank
A	BIO 9900	0.42	9	0.65	10	0.28	5	0.14	2
B	BIO 1263	0.34	8	0.48	8	0.70	11	0.64	9
C	BIO 1169	0.28	6	0.66	11	0.27	4	0.54	8
D	BC 42521	0.78	11	0.11	1	0.62	9	0.72	12
E	BC 42683	0.12	2	0.24	5	0.23	3	0.20	4
F	BC 41399	0.10	1	0.19	3	0.15	1	0.12	1
G	BC 2630	0.28	5	0.42	6	0.67	10	0.71	10
H	BC 42882 –A	0.32	7	0.19	2	0.46	8	0.39	7
I	BIO 9899	0.23	3	0.22	4	0.16	2	0.32	5
J	BISI – 2	0.78	12	0.50	9	0.86	12	0.72	11
K	P – 12	0.44	10	0.70	12	0.44	7	0.19	3
L	C 7	0.25	4	0.47	7	0.33	6	0.38	6

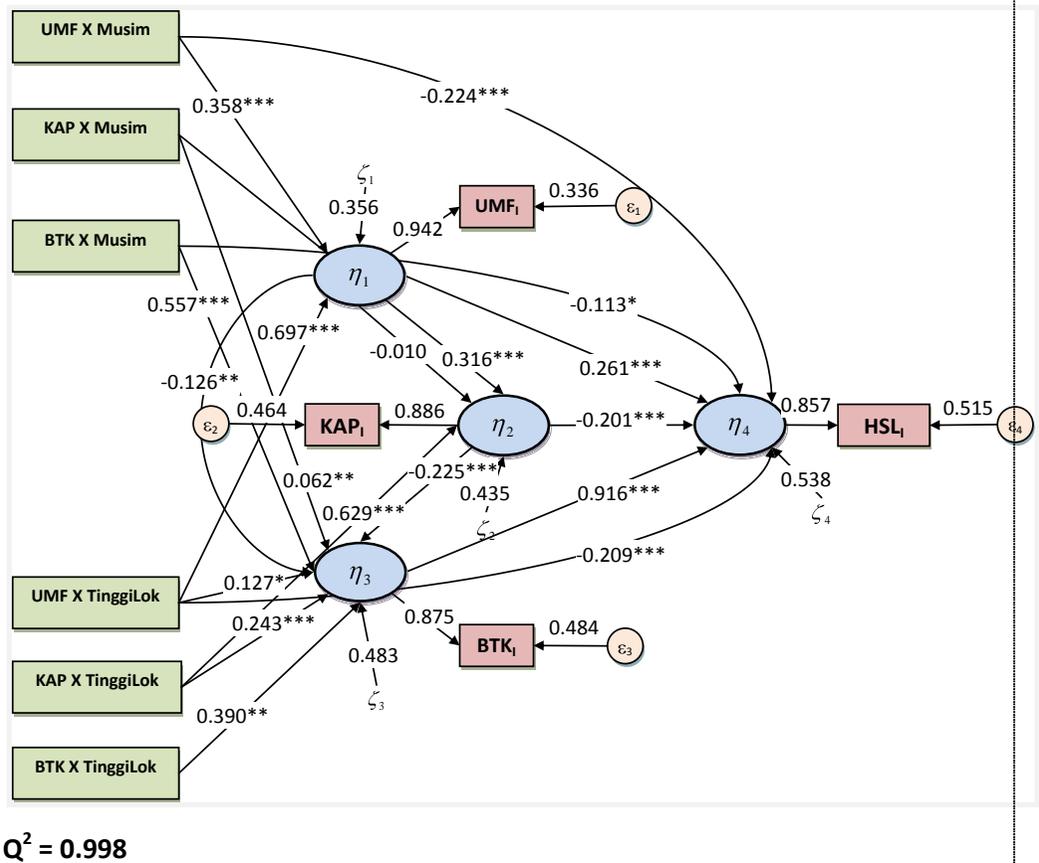
Tabel 6. Hasil Klasifikasi Genotipe

Lingkungan		Respon			
		UMF	KAP	BTK	DH
		Genotipe Spesifik Lokasi			
L1	Ketaon		E,L	D,K	
L2	Kemiri	C	G		D,C
L3	Moncongloe Bulu			B	I,L,J
L4	Yoso Mulyo-H	G,B,K	I,J	C,A	D,C
L5	Sido waras	G,B,K		L,J	B
L6	Brodot		A		E
L7	Wringinsongo		A		H,G
L8	Kuta Tengah-H	H, J	I,J	H,G	E,H,G
L9	Cempedak Lobang	H, J			H
L10	Pabuaran	D,C	B,K		
L11	Kalikotes			C,A	C,D
L12	Kuta Tengah-K	L	G,C		C,D
L13	Sambirejo	L	G,C		
L14	Yoso Mulyo-K	G,B,K		C,A	C,D
L15	Pontang	A, L			,D
L16	Jambu Timur	A, L	L, E	L,J	I,L,J
		Genotipe Stabil			
		F,E,I	D, H,F	F,I,E	F,A,K

Partial Least Square – AMMI

Pendekatan metode Partial Least Square (PLS)-AMMI merupakan penggabungan konsep PLS dengan metode AMMI. AMMI digunakan untuk mendapatkan matriks interaksi sebagai skor laten interaksi sedangkan PLS digunakan untuk memodelkan matriks interaksi tersebut. PLS- AMMI digunakan dalam uji multilokasi memiliki keuntungan dalam menjelaskan sumbangan dari komponen-komponen daya hasil dan faktor lokasi. Metode PLS digunakan untuk uji stabilitas genotipe pertama kali diperkenalkan oleh Aastveit H pada tahun 1986 namun hanya memperhatikan faktor lingkungan belum memasukkan informasi komponen daya hasil. Metode ini dikembangkan dengan mengadopsi pemikiran Aastveit, (1986) dan Dhungana (2004).

Hasil pendugaan parameter model struktural dengan PLS disajikan di bawah ini.



Gambar 6. Diagram Jalur IGL Daya Hasil

Model dasar dari penelitian ini ditunjukkan pada Gambar 1. Pemilihan variabel eksogen yaitu kombinasi kovariat genotipik \times lingkungan yang dimasukkan ke dalam model dilakukan dengan prosedur maju (*forward selection procedure*) yaitu memasukkan satu-persatu variabel eksogen ke dalam model kemudian dilakukan evaluasi kecocokan model. Jika dengan memasukkan variabel eksogen tertentu variabel tersebut memberikan pengaruh secara nyata dan model cocok, maka variabel tersebut dipertahankan dalam model jika tidak variabel tersebut dikeluarkan dari model.

Hasil analisis PLSPM diketahui bahwa nyatanya efek IGL daya hasil pada percobaan multilokasi dipengaruhi oleh IGL komponen-komponen daya hasil yaitu IGL usia masak fisiologis, kadar air panen, dan berat tongkol panen dengan *standardized* efek masing-masing adalah 0.331, -0.204, 0.921 seperti yang terlihat pada Gambar 6. Terlihat berat tongkol panen memberikan efek langsung terbesar. Secara pengaruh total, berat tongkol panen memberikan efek terbesar pada IGL daya hasil. Hasil ini menunjukkan bahwa berat tongkol panen merupakan indikator stabilitas utama dari Daya Hasil. Hasil analisis PLSPM-AMMI mendukung bahwa BC 41399, BC BIO 9899 dan BC 42683 adalah genotipe-genotipe relative stabil dari kajian struktur interaksi komponen IGL daya hasil dan IGL daya hasil karena komponen-komponen daya hasil merupakan indikator stabilitas daya hasil. IGL komponen daya hasil usia masak fisiologis memiliki efek langsung positif terhadap IGL daya hasil. Ini artinya bahwa untuk usia masak fisiologis di atas rata-rata maka memiliki daya hasil yang relatif lebih tinggi. Begitu juga IGL berat tongkol. Namun, untuk IGL kadar air panen memiliki efek langsung negatif terhadap IGL daya hasil. Hasil ini memberikan informasi bahwa jika kadar air panen di atas rata-rata maka daya hasil atau hasil produksi jagung relatif lebih sedikit. Ini mungkin terjadi karena adanya proses pengeringan dimana daya hasil dihitung untuk kadar air 15%. IGL usia masak fisiologis yang memberikan efek tidak langsung melalui kadar air panen, dan berat tongkol sebesar -0.090. Tanda negatif ini terjadi karena melalui kadar air panen yang memiliki efek negatif pada daya hasil. Selanjutnya, kadar air panen juga memberikan efek tidak langsung terhadap IGL daya hasil melalui berat tongkol dengan besar efek tidak langsungnya adalah -0.209. Total efek dari ketiga IGL komponen daya hasil secara berurutan adalah 0.241 dari IGL usia masak fisiologis, -0.413 dari IGL kadar air panen, dan 0.921 dari IGL berat tongkol. Dari model PLSPM ini juga dapat diketahui keragaman dari IGL usia masak fisiologis, IGL kadar air panen, dan IGL berat tongkol panen dan IGL daya hasil yang dapat dijelaskan oleh model secara berurutan adalah 0.886, 0.816, 0.763 dan 0.721 dengan keragaman total dihitung dari nilai Q^2 adalah sebesar 0.999. Besarnya

nilai-nilai ini menunjukkan bahwa model yang dianalisis dapat menjelaskan keterkaitan antara IGL komponen daya hasil, pengaruhnya terhadap daya hasil dan mampu menjelaskan pengaruh kombinasi kovariat genotipik x lingkungan terhadap IGL Daya Hasil.

Kovariat genotipik x lingkungan yang berpengaruh nyata terhadap daya hasil pada taraf nyata 5% adalah usia masak fisiologi x tinggi lokasi, usia masak fisiologis x musim, dan berat tongkol x musim. Hasil ini memberikan gambaran bahwa jagung dengan usia masak fisiologis di atas rata-rata ditanam pada lokasi relatif tinggi maka akan berakibat pada daya hasil yang kurang baik. Begitu juga jika di tanam pada musim dimana hujannya relatif tinggi. Sedangkan jika untuk genotipe dengan berat tongkol panen di atas rata-rata jika di tanam pada lokasi dengan tingkat curah hujan yang relatif tinggi maka daya hasil jagung akan relatif kurang baik.

Hasil dari PLS-AMMI dapat digunakan sebagai identifikasi awal genotipe unggulan dengan beberapa kriteria yaitu stabilitas relatif tinggi, daya hasil di atas rata-rata, usia masak fisiologis di atas rata-rata, kadar air panen di bawah rata-rata, dan berat tongkol di atas rata-rata. Faktor lingkungan yang relatif baik adalah curah hujan relatif rendah dan lokasi di dataran rendah. Dari kriteria di atas maka genotipe BC 42683 dan BC 41399 dan BIO 9899 dapat dipertahankan untuk uji lebih lanjut dalam seleksi genotipe unggulan.

Untuk peningkatan produksi jagung, dari analisis PLSPM-AMMI diperoleh prioritas komponen daya hasil yang harus diperhatikan secara berurutan adalah berat tongkol panen, kadar air panen dan usia masak fisiologis.

5. KESIMPULAN DAN SARAN

Kajian struktur interaksi komponen-komponen daya hasil dan struktur interaksi daya hasil dengan metode AMMI menemukan tiga genotipe yang relatif stabil yaitu BC 41399, BC BIO 9899 dan BC 42683.

Melalui PLSPM-AMMI dapat dilakukan pemodelan IGL Daya Hasil dengan memperhatikan sumbangan dari komponen daya hasil dan kombinasi kovariat genotipik x lingkungan yang berguna untuk mengidentifikasi faktor-faktor yang menyebabkan nyatanya efek interaksi genotipe x lokasi dalam percobaan multilokasi Serta menemukan komponen daya hasil dan kombinasi genotipik x lingkungan yang paling berperan dalam peningkatan hasil produksi.

Hasil analisis PLS-AMMI menunjukkan bahwa indikator utama stabilitas dari daya hasil adalah berat tongkol panen, kemudian kadar air panen dan terakhir usia

masak fisiologis. Sehingga proses seleksi genotipe harus memperhatikan ketiga komponen tersebut sesuai urutan prioritasnya. Kombinasi kovariat genotipik x lingkungan yang berpengaruh nyata terhadap daya hasil adalah usia masak fisiologis x musim, usia masak fisiologis x tinggi dan berat tongkol panen x musim. Hasil analisis PLS-AMMI juga mengidentifikasi genotipe BC 41399, BC BIO 9899 dan BC 42683 untuk dipertimbangkan sebagai genotipe unggulan

6. DAFTAR PUSTAKA

- Aastveit H., and H. Martens. 1986. ANOVA Interactions Interpreted by Partial Least Squares Regression, *Biometrics* 42 : 829-844
- Bollen, K.A. 1989. *Structural Equation With Latent Variabels*. John Wiley and Sons, New York
- Dhungana, P. 2004. Partial least square path modeling of Genotype x Environment Interaction. Ph. D. Dissertation. University of Nebraska-Lincoln, Lincoln, Nebraska.
- Gauch, H.G. JR. 1982. Noise reduction by eigenvector ordination. *Ecology* 63:1643-1649
- Jeffrey J. Hoogland. 1971. *The Robustness of Estimation Methods for Covariance Structure Analysis*. Thela Thesism, Groningen
- Nur et al. 2007. Stabilitas Komponen Hasil sebagai Indikator Stabilitas Hasil Genotipe Jagung Hibrida, *Penelitian Pertanian Tanaman Pangan* Vol. 26 No. 2. Balai Penelitian Tanaman Serealia
- Mattjik, AA & Sumertajaya IM 2000. *Perancangan Percobaan dengan Aplikasi SAS dan Minitab* jilid 1. IPB Press, Bogor.

Halaman ini sengaja dikosongkan

Pengamatan Pencilan pada Analisis Interksi Genotipe Lingkungan

8

PENGAMATAN PENCILAN PADA ANALISIS INTERKSI GENOTIPE LINGKUNGAN

Pengamatan pencilan sering mendapat perhatian pada analisis statistika. Ukuran pemusatan dan penyebaran data yang sering digunakan seperti rata-rata mempunyai sifat rentan terhadap adanya pencilan. Akibat satu pengamatan yang sangat jauh berbeda nilainya dari sesamanya akan cenderung mempengaruhi nilai rata-rata. Demikian juga dengan ragam, karena ragam diukur menggunakan konsep rata-rata jarak kuadrat pengamatan terhadap rata-ratanya maka ia juga rentan terhadap pencilan. Hampir semua prosedur statistik yang mendasarkan pada statistik rata-rata dan ragam dan sebaran normal pada umumnya, secara teori akan menghadapi hal yang sama. Termasuk di dalamnya pendekatan kuadrat (jarak) terkecil atau Least Square.

Pencilan (pada satu arah tertentu) memungkinkan menjadi pengaruh dalam pengujian ketaknormalan. Dalam hal ini diperlukan kajian sensitifitas pengaruh pencilan terhadap pengujian ketaknormalan. Mendeteksi suatu pengamatan

sebagai pencilan dari suatu sebaran normal atau sebaran simetrik, atukah pengamatan tersebut menjadi bagian dari suatu sebaran tertentu yang tidak simetrik. Model AMMI (Additive Main Effect Multiplicative Interaction) memerlukan asumsi parametrik tertentu yang harus dipenuhi. AMMI merupakan metode yang menggabungkan analisis ragam pada pengaruh aditif dengan analisis komponen utama pada pengaruh multiplikatif.

8.1 Stabilitas Non Parametrik

Kajian Interaksi genotipe dan lingkungan (IGL) yang dilakukan oleh pemulia tanaman sangat penting dilakukan. Hasilnya yang berupa klasifikasi genotipe diharapkan dapat digunakan untuk mengidentifikasi genotipe yang berpenampilan stabil pada berbagai lingkungan berbeda atau beradaptasi pada suatu lingkungan spesifik. Genotipe stabil dengan hasil tinggi (genotipe stabil dinamik) lebih diperhatikan oleh pemulia tanaman karena kebutuhan pasar dan lamanya waktu yang diperlukan untuk mengembangkan genotipe.

Metode AMMI dimulai dengan melihat pengaruh aditif genotipe dan lingkungan dengan menggunakan analisis ragam dan kemudian dibuat bentuk multiplikatif interaksi genotipe dengan lingkungan menggunakan analisis komponen utama. Bentuk multiplikatif diperoleh dari penguraian interaksi genotipe dengan lingkungan menjadi komponen utama interaksi (KUI) menggunakan metode penguraian nilai singular (Singular Value Decomposition, SVD). Alat yang digunakan untuk menginterpretasi hasil metode AMMI adalah biplot AMMI. Analisis kestabilan genotipe pada AMMI dapat ditentukan dengan menggunakan indeks stabilitas AMMI (ISA). Penggunaan indeks stabilitas dilakukan untuk menentukan peringkat stabilitas suatu genotipe terhadap lingkungan. Indeks dibangun berdasarkan konsep jarak sehingga semakin kecil indeks suatu genotipe maka semakin dekat jarak genotipe dari pusat sumbu koordinat, artinya semakin stabil genotipe tersebut. ISA menggunakan m buah KUI yang nyata.

Model AMMI dibangun dengan landasan teori pemodelan parametrik pada distribusi data Normal (gaussian). Analisis menggunakan pendekatan parametrik memerlukan asumsi tertentu yang harus dipenuhi seperti asumsi kenormalan dan kehomogenan ragam pada galat. Hasil analisis yang baik akan diperoleh hanya jika asumsi-asumsi parametrik terpenuhi. Selain itu AMMI secara teori, AMMI diperkirakan juga memiliki kerentanan terhadap adanya pengamatan pencilan. Hal ini dapat dijelaskan karena pada pengaruh multiplikatif, AMMI menguraikan struktur interaksi dari matriks sisaan komponen aditif dengan

memanfaatkan sifat matematis penguraian nilai singular (singular value decomposition, SVD) yang berbasis Least Square. Pada data numerik yang mengikuti sebaran normal, keberadaan pencilan pada satu sisi tertentu akan menimbulkan kemenjuluran (skewness) yang cenderung terdeteksi sebagai ketaknormalan. Sehingga pencilan dapat pula menjadi penyebab tidak terpenuhinya asumsi dasar pendekatan parametrik.

Sedangkan analisis kestabilan menggunakan pendekatan nonparametrik tidak memerlukan asumsi sebaran tertentu dan dilakukan berdasarkan peringkat. Beberapa pendekatan nonparametrik untuk memperoleh klasifikasi genotipe telah banyak dilakukan di antaranya metode Huehn, metode Kang, metode Fox, dan metode Thennarasu. Huehn dalam Akcura (2008) menyatakan bahwa mengukur stabilitas berdasarkan peringkat mempunyai kelebihan dibandingkan pendekatan parametrik, yaitu mereduksi bias yang berasal dari pencilan, tidak adanya asumsi yang diperlukan pada nilai pengamatan serta mudah digunakan dan diinterpretasikan. Berbeda dengan pengukuran parametrik yang sensitif terhadap asumsi statistika yang harus dipenuhi, pendekatan nonparametrik dapat menjadi alternatif pengukuran yang kekar (robust) jika pelanggaran asumsi parametrik terjadi.

Zulhayana 2010 melakukan perbandingan penggunaan beberapa Indeks Stabilitas Nonparametrik (ISN) yaitu metode Huehn, metode Kang, metode Fox, dan metode Thennarasu. Indeks Top pada metode Fox adalah ISN terbaik diantara ISN lainnya karena memiliki korelasi tertinggi dengan hasil. Indeks Top juga memiliki kemiripan hasil klasifikasi genotipe dengan ISA. Evaluasi terhadap kekekaran ISN Top (metode Fox), Zulhayana 2010 menyebutkan bahwa, pada kasus pencilan yang terjadi pada pengamatan ulangan tertentu dari Genotipe \times Lingkungan tertentu, ISN Top tidak menunjukkan sifat kekar, namun pada kasus nilai rata-rata Genotipe \times Lingkungan tertentu yang ekstrim ISN Top memiliki sifat kekar. Meski tidak menunjukkan sifat kekekaran terhadap adanya pencilan pada pengamatan ulangan, setidaknya penggunaan ISN harus tetap dipertimbangkan pada saat beberapa asumsi parametrik yang tidak terpenuhi. Sedangkan pada kasus adanya nilai rata-rata IGL yang ekstrim maka penggunaan ISN dengan sifat kekarnya akan sangat membantu.

8.2 Robust-AMMI dengan Model Faktor Analitik

Sebagaimana telah disebutkan sebelumnya bahwa Model AMMI dan model-model lain yang menggunakan Singular Value Dekomposition (SVD) seperti Analisis Komponen Utama dan Analisis Faktor, rentan terhadap adanya pencilan. Sedangkan dalam upaya merakit sifat-sifat unggul pada suatu genotipe, pencilan

justru menjadi sesuatu yang berharga, karenanya mengabaikannya tidaklah bijaksana. Untuk itu diperlukan metode yang relatif kekar terhadap adanya pencilan.

Selain rataan dan ragam, dikenal pula ukuran pemusatan dan penyebaran lain yaitu median dan simpangan mutlak median atau *Median Absolute Deviation* (MAD). Dua statistik ini berbasis median yang relatif kekar terhadap pencilan. Beberapa prosedur statistika juga dikebangkan menurut keduanya, antara lain regresi kekar dengan kosep simpangan median mutlak terkecil (*least absolut*) menggantikan kuadrat terkecil (*least square*). Prosedur-prosedur ini kemudian disebut dalam kelas pemodelan atau pendekatan *robust*.

Dalam kasus penerapan model AMMI untuk analisis IGL dan kestabilan, meski belum pernah dilaporkan sebelumnya, namun karena secara teoritik SVD berbasis kuadrat terkecil, maka tetap berpotensi menghadapi masalah akibat pengamatan pencilan. Ide dasar membangun kekekaran pada model AMMI maka dasarnya adalah mengubah kriteria pemodelan dari *least square* ke *least absolute*, baik itu pada pendugaan pengaruh utama dan pada pengaruh interaksi. Cukup sulit untuk membayangkan bagaimana analisis ragam yang sudah begitu melekat dalam kehidupan ilmiah harus kita ubah menjadi analisis simpangan mutlak, atau istilah lain bagi pedekatan robust. Betapapun sulit dan tidak populer, hal tersebut tetap harus dilakukan sebagai suatu langkah ilmiah yang visioner.

Mungkin jalan ke arah itu masih jauh. Namun pada awal sejarah perkembangan AMMI menunjukkan bahwa terdapat benang merah antara model AMMI dan model Faktor pada pemodelan aditif-multiplikatif, yaitu dalam dekomposisi suku multiplikatif. Gollob 1968 mengenalkan model faktor-analisis ragam (disebut FANOVA) yang menggabungkan aspek analisis ragam dan faktor analisis. Sedangkan Gabriel 1978 mendiskusikan beberapa model yang mirip dengan FANOVA dengan pendugaan parameter menggunakan metode kudrat terkecil. Model AMMI kemudian berkembag dan dikenal sesudah itu. Cornelius, et al. 1996 menuliskan bahwa Gaugh & Zobel memberi nama lain pada model FANOVA yang diusulkan Gollob sebagai model AMMI. Perkembangan terakhir dalam upaya mengatasi ketakhomogenan pada AMMI dalam konteks model campuran, Smith et al., 2002 mengusulkan model yang disebut faktor analitik multiplikatif. Karena menurutnya model faktor analitik ekuivalen dengan AMMI model campuran. Seiring dengan perkembangan dekomposisi matriks telah pula dikenal "robust" *Principal Component Analysis* (Jolliffe, 1986), dan "robust" *Factor Analysis* (Croux & Filzmoser, 1998) yang relatif kekar terhadap

pengamatan pencilan, maka dalam kelas model faktor analitik ini berkembang pula model faktor analitik yang kekar.

Pengembangan model robust berawal dari penggunaan model faktor dan menggunakan pendekatan regresi bolak balik dengan kriteria *least absolute*. Suatu algoritma Regresi Bolak-balik diusulkan oleh Croux *et al.* (2003) digunakan untuk mengkonstruksi *Robust Factor Analysis*. Pada dasarnya ide membangun model Robus FANOVA adalah dengan terlebih dahulu menduga efek aditif dengan model kekar terlebih dahulu, dan secara berturutan kemudian melakukan analisis faktor yang kekar (*robust factor*) pada residual. Tapi cara ini hanya berlaku bagi pengepasan model dengan pengekatan kuadrat terkecil dan tidak lagi berlaku bagi model FANOVA yang kekar. Oleh karena itu kita akan menduga suku aditif dan multiplikatif secara simultan dengan menerapkan *Robust Alternating Regression (RAR)* dalam model FANOVA, melalui penggunaan *least absolute (L1 Regression)*. Model RAR-FANOVA menurunkan bobot (*downweighting*) bagi pencilan skor dan *loadings* di ruang berdimensi- k , pembobotan pada regresi L1 ini dimaksudkan untuk menurunkan pengaruh pengamatan pengungkit (*leverage points*). Kemudian secara unik mengidentifikasi parameter dalam pengamatan. Fungsi penduga RAR diminimalkan bawah kendala yang konsisten dengan pendekatan robus bagi parameter median. Ketika tidak ada pencilan yang sangat parah dan mengganggu, maka metode robust memberikan hasil sedikit saja berbeda dari pendekatan kuadrat terkecil.

Pada bagian berikutnya disajikan dua artikel tentang Stabilitas NonParametrik dan model Robust.

Halaman ini sengaja dikosongkan

8.3 Analisis Stabilitas Genotipe Padi dengan Indeks Stabilitas Nonparametrik Thennarasu (S. Zulhayana, I M Sumertajaya, A A Mattjik)

Analisis Stabilitas Genotipe Padi dengan Indeks Stabilitas Nonparametrik Thennarasu

Sri Zulhayana^{1,2}, I M Sumertajaya³, A A Mattjik³

¹Mahasiswa Pascasarjana Departemen Statistika, Institut Pertanian Bogor

²Jurusan Pendidikan Matematika, FMIPA STKIP Budidaya Binjai

³Departemen Statistika, FMIPA Institut Pertanian Bogor

Abstrak. Pemulia tanaman seringkali mengalami kesulitan dalam seleksi genotipe unggulan pada percobaan multilokasi. Analisis stabilitas diperlukan untuk dapat mengidentifikasi genotipe dengan penampilan stabil dinamik pada berbagai lingkungan berbeda. Terdapat pendekatan parametrik dan nonparametrik dalam menganalisis stabilitas genotipe. Analisis stabilitas nonparametrik dilakukan berdasarkan peringkat genotipe di tiap lingkungan dan tidak memerlukan asumsi distribusi tertentu. Salah satu diantara analisis stabilitas nonparametrik yaitu menggunakan indeks stabilitas nonparametrik (ISN) Thennarasu yang terdiri atas 4 indeks $NP_i^{(1)}$, $NP_i^{(2)}$, $NP_i^{(3)}$ dan $NP_i^{(4)}$. Penelitian ini dilakukan untuk memperoleh klasifikasi genotipe menggunakan ISN Thennarasu dan membandingkan ISN Thennarasu dari 14 genotipe padi di 20 lingkungan. Hasil penelitian menunjukkan genotipe L memiliki indeks terkecil pada keempat ISN Thennarasu. Oleh karena itu genotipe L diidentifikasi sebagai genotipe yang paling stabil. Adapun genotipe B diidentifikasi sebagai genotipe yang paling tidak stabil. Berdasarkan korelasi peringkat rata-rata hasil dengan peringkat ISN Thennarasu diperoleh bahwa $NP_i^{(3)}$ dan $NP_i^{(4)}$ memiliki korelasi tertinggi dengan rata-rata hasil yaitu sebesar -0.943.

Kata kunci: indeks stabilitas nonparametrik Thennarasu, genotipe stabil

PENDAHULUAN

Kajian Interaksi genotipe dan lingkungan (IGL) yang dilakukan oleh pemulia tanaman sangat penting dilakukan. Hasilnya yang berupa klasifikasi genotipe

diharapkan dapat digunakan untuk mengidentifikasi genotipe yang berpenampilan stabil pada berbagai lingkungan berbeda atau beradaptasi pada suatu lingkungan spesifik. Genotipe stabil dengan hasil tinggi (genotipe stabil dinamik) lebih diperhatikan oleh pemulia tanaman karena kebutuhan pasar dan lamanya waktu yang diperlukan untuk mengembangkan genotipe.

Terdapat dua pendekatan dalam mengkaji interaksi genotipe dan lingkungan yaitu pendekatan parametrik dan nonparametrik. Analisis menggunakan pendekatan parametrik memerlukan asumsi tertentu yang harus dipenuhi seperti asumsi kenormalan dan kehomogenan ragam pada galat. Hasil analisis yang baik akan diperoleh selama asumsi parametrik terpenuhi. Sedangkan analisis menggunakan pendekatan nonparametrik tidak memerlukan asumsi sebaran tertentu dan dilakukan berdasarkan peringkat. Beberapa pendekatan nonparametrik untuk memperoleh klasifikasi genotipe telah banyak dilakukan di antaranya metode Huehn, metode Kang, metode Fox, dan metode Thennarasu.

Huehn dalam Akcura (2008) menyatakan bahwa mengukur stabilitas berdasarkan peringkat mempunyai kelebihan dibandingkan pendekatan parametrik, yaitu mereduksi bias yang berasal dari pencilan, tidak adanya asumsi yang diperlukan pada nilai pengamatan serta mudah digunakan dan diinterpretasikan. Berbeda dengan pengukuran parametrik yang sensitif terhadap asumsi statistika yang harus dipenuhi, pendekatan nonparametrik dapat menjadi alternatif pengukuran yang kekar (robust) jika pelanggaran asumsi parametrik terjadi. Dalam penelitian ini akan dilakukan pengkajian klasifikasi genotipe yang dihasilkan oleh pendekatan nonparametrik Thennarasu.

DATA DAN METODE

Data

Penelitian ini melibatkan 14 genotipe padi di 20 lingkungan, yang diseleksi dari Balai Besar Penelitian Padi. Percobaan dilakukan pada tahun 2008 dengan tujuan mengevaluasi penampilan fenotipe dari padi dari generasi terakhir dalam lingkungan berbeda. Percobaan dilakukan menggunakan rancangan acak kelompok lengkap dengan 3 ulangan dalam tiap lingkungan. Ukuran tiap plot adalah 4 m x 5 m dengan jarak 25 cm. Pupuk yang digunakan yaitu 250 kg N ha⁻¹, 100 kg P ha⁻¹ dan 100 kg K ha⁻¹. Rataan hasil diduga untuk tiap genotipe di tiap lingkungan.

Metode

Pada penelitian ini digunakan ISN Thennarasu sebagai acuan pembentukan klasifikasi genotipe. Indeks stabilitas dengan metode Thennarasu dihitung berdasarkan peringkat r_{ij} dan r_{ij}^* ; r_{ij} adalah peringkat data genotipe ke- i dan

lingkungan ke-j sedangkan r_{ij}^* adalah peringkat data yang terkoreksi dengan rata-rata genotipe tiap lingkungan. Terdapat empat macam ISN pada metode Thennarasu sebagai berikut:

$$NP_i^{(1)} = \frac{1}{l} \sum_{j=1}^l |r_{ij}^* - M_{di}^*|$$

$$NP_i^{(2)} = \frac{1}{l} \left(\sum_{j=1}^l |r_{ij}^* - M_{di}^*| / M_{di} \right)$$

$$NP_i^{(3)} = \frac{\sqrt{\sum (r_{ij}^* - \bar{r}_i^*)^2 / l}}{\bar{r}_i}$$

$$NP_i^{(4)} = \frac{2}{l(l-1)} \left[\sum_{j=1}^{l-1} \sum_{(j=j+1)}^l |r_{ij}^* - r_{ij}^*| / \bar{r}_i \right]$$

Dalam ISN Thennarasu tersebut, r_{ij}^* merupakan peringkat x_{ij}^{**} ; dengan x_{ij}^{**} merupakan rata-rata data yang terkoreksi dengan rata-rata genotipe tiap lingkungan yaitu $x_{ij}^{**} = x_{ij} - \bar{x}_i$; Sedangkan \bar{r}_i^* dan M_{di}^* adalah rata-rata dan median r_{ij}^* . Adapun \bar{r}_i dan M_{di} adalah rata-rata dan median r_{ij} ; dengan r_{ij} adalah peringkat data asli. Klasifikasi genotipe dengan metode Thennarasu ditentukan berdasarkan ISN yang diperoleh. Semakin besar ISN genotipe maka semakin stabil genotipe tersebut.

ANALISIS STABILITAS

Hasil perhitungan ISN Thennarasu dirangkum pada Tabel. Nilai ISN yang lebih kecil mengindikasikan genotipe lebih stabil dibandingkan dengan yang lain. Berdasarkan ISN $NP_i(1)$, $NP_i(2)$, $NP_i(3)$ dan $NP_i(4)$ pada Tabel 1 dapat dilihat bahwa genotipe L memiliki ISN terkecil.

Tabel 1 Indeks stabilitas nonparametrik Thennarasu

Genotipe	$NP_i^{(1)}$	$NP_i^{(2)}$	$NP_i^{(3)}$	$NP_i^{(4)}$
A	3.61	0.89	0.96	1.16
B	4.50	1.47	1.32	1.55
C	2.77	0.39	0.45	0.53
D	3.02	0.26	0.39	0.45
E	3.34	0.30	0.41	0.48

Model AMMI Kini dan Yang Akan Datang

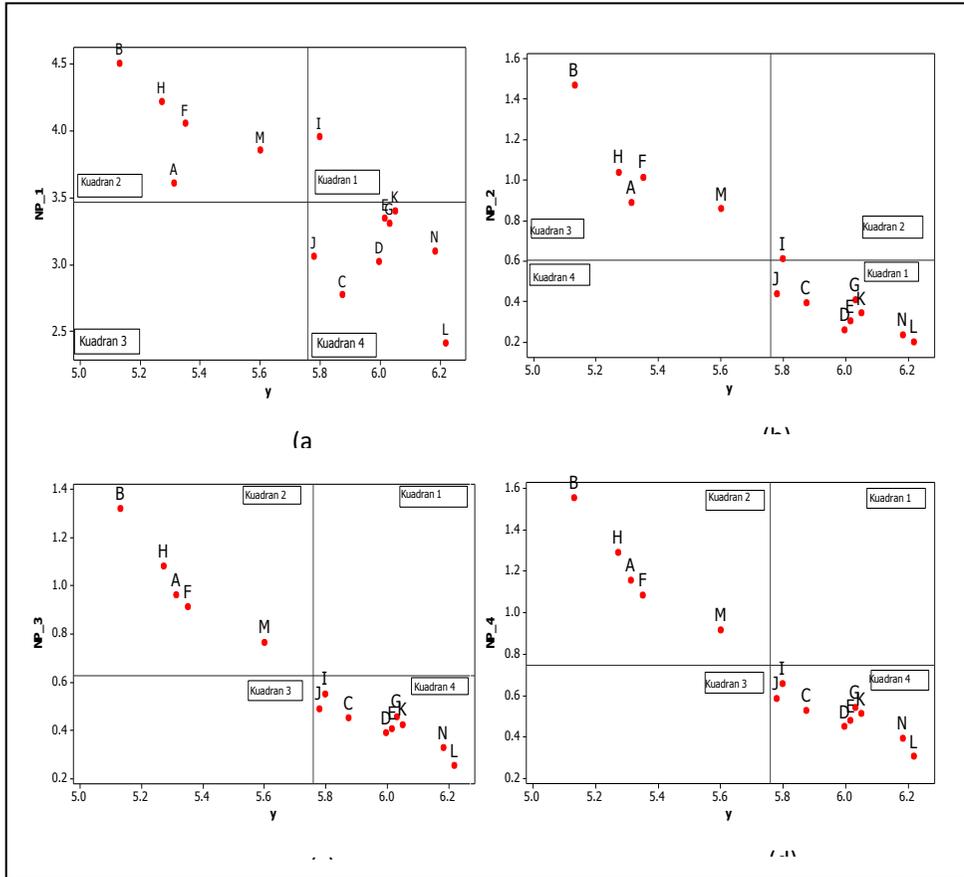
F	4.05	1.01	0.91	1.08
G	3.31	0.41	0.45	0.54
H	4.22	1.04	1.08	1.29
I	3.95	0.61	0.55	0.66
J	3.06	0.44	0.49	0.58
K	3.40	0.34	0.42	0.51
L	2.41	0.20	0.25	0.30
M	3.85	0.86	0.76	0.91
N	3.10	0.23	0.32	0.39

Tabel 2 Klasifikasi genotipe dengan ISN Thennarasu

Peringkat stabilitas	Genotipe			
	$NP_i^{(1)}$	$NP_i^{(2)}$	$NP_i^{(3)}$	$NP_i^{(4)}$
1	L	L	L	L
2	C	N	N	N
3	D	D	D	D
4	J	E	E	E
5	N	K	K	K
6	G	C	C	C
7	E	G	G	G
8	K	J	J	J
9	A	I	I	I
10	M	M	M	M
11	I	A	F	F
12	F	F	A	A
13	H	H	H	H
14	B	B	B	B

Karena itu, keempat ISN Thennarasu tersebut mengidentifikasi genotipe L sebagai genotipe yang paling stabil. Sedangkan genotipe dengan ISN terbesar dimiliki oleh genotipe B. Jadi genotipe B diidentifikasi sebagai genotipe yang paling tidak stabil. Selanjutnya, urutan selengkapnya genotipe stabil berdasarkan ISN $NP_i^{(1)}$, $NP_i^{(2)}$, $NP_i^{(3)}$ dan $NP_i^{(4)}$ dapat dilihat pada Tabel 2. Tabel 2 menunjukkan bahwa hasil klasifikasi genotipe berdasarkan ISN $NP_i^{(1)}$ dan $NP_i^{(2)}$ adalah sama. Kedua ISN ini serupa dalam mengidentifikasi stabilitas keempatbelas genotipe stabil. Sedangkan $NP_i^{(1)}$ dan $NP_i^{(2)}$ menghasilkan klasifikasi genotipe yang tidak sama.

Model AMMI Kini dan Yang Akan Datang



Gambar 1 (a) Plot rata-rata hasil dengan $NP_i^{(1)}$; (b) Plot rata-rata hasil dengan $NP_i^{(2)}$
 (c) Plot rata-rata hasil dengan $NP_i^{(3)}$; (d) Plot rata-rata hasil dengan $NP_i^{(4)}$

ISN Thennarasu terbaik

Tiap ISN Thennarasu menghasilkan peringkat genotipe yang unik. ISN Thennarasu terbaik dapat dilihat melalui korelasi peringkat Spearman ISN Thennarasu dengan rata-rata hasil (Y). Berdasarkan hasil perhitungan korelasi (Tabel 3), keempat ISN Thennarasu memiliki korelasi negatif dengan hasil. Negatifnya nilai korelasi ini dikarenakan kriteria ISN Thennarasu dimana genotipe paling stabil adalah genotipe yang memiliki ISN Thennarasu terkecil dan hasil terbesar.

Tabel 3 Korelasi Peringkat Spearman antara hasil dan peringkat ISN Thennarasu

Indeks	Y	$NP^{(1)}$	$NP^{(2)}$	$NP^{(3)}$
$NP^{(1)}$	-0.736			
$NP^{(2)}$	-0.938	0.851		

Model AMMI Kini dan Yang Akan Datang

NP ⁽³⁾	-0.943	0.837	0.996	
NP ⁽⁴⁾	-0.943	0.837	0.996	1.00

Korelasi tertinggi dengan rata-rata hasil dimiliki oleh NPi(3) dan NPi(4). Sehingga NPi(3) dan NPi(4) merupakan ISN Thennarasu terbaik dalam penelitian ini. Selain itu, NPi(3) dan NPi(4) memiliki korelasi paling kuat yaitu 1. Karena itu, dapat dipilih satu dari NPi(3) dan NPi(4) sebagai ISN yang dipakai untuk mengklasifikasi genotipe stabil.

KESIMPULAN

Keempat ISN Thennarasu mengindikasikan genotipe L sebagai genotipe yang paling stabil dan genotipe B sebagai genotipe yang paling tidak stabil. NPi(3) dan NPi(4) merupakan ISN Thennarasu terbaik dalam penelitian ini dan dapat dipilih satu dari kedua ISN tersebut sebagai rekomendasi untuk mengklasifikasi genotipe stabil.

DAFTAR PUSTAKA

- Akcura M, Kaya Y. 2008. *Nonparametric Stability Methods for Interpreting Genotype by Environment Interaction of Bread Wheat Genotypes (Triticum aestivum L.)*. *Genetics and Molecular Biology* 31: 906-913.
- Kang MS. 2002. *Genotype-Environment Interaction: Progress and Prospects*. Di dalam : Kang MS, Editor. *Quantitative Genetics, Genomic and Plant Breeding*. Florida: CRC Pr. Hlm. 221-243.
- Kaya Y, Taner S, Ceri S. 2003. Nonparametric Stability Analysis of Yield Performances in Oat Genotypes across Environments. *Asian Journal of Plant Sciences*. 2(3): 286-289.
- Mattjik AA, Sumertajaya IM. 2006. *Perancangan Percobaan dengan Aplikasi SAS dan Minitab* Jilid I Edisi 2. Bogor: IPB Press.
- Mohammadi R, Abdulahi A, Haghparast R, Aghaee M, Rostaee M. 2007. Nonparametric Methods for Evaluating of Winter Genotypes in Multi-environment Trials. *World Journal of Agricultural Sciences* 3(2):137-242
- Sabaghnia N, Dehghani H, Sabaghpour SH. 2006. Nonparametric Methods for Interpreting Genotype x Environment Interaction of Lentil Genotypes. *Crop Sci* 46: 1100-1106.

Solomon KF, Smith HA, Malan E, Du Toit WJ. 2007. Comparison Study Using Rank Based Nonparametric Stability Statistics of Durum Wheat. *World Journal of Agricultural Sciences* 3(4): 444-450.

Halaman ini sengaja dikosongkan

8.4 Handling Outlier Observation in Two-Ways Table by Robust Alternating Regression of Fanova Models: Towards Robust AMMI Models (Alfian F. Hadi & A. A. Mattjik, Jurnal ILMU DASAR V. 12 No. 2, 2011, in press)

Handling Outlier Observation in Two-Ways Table by Robust Alternating Regression of Fanova Models: Towards Robust AMMI Models

Alfian F. Hadi¹ and A. A. Mattjik²

¹ PhD Student. Graduate School of Bogor University of Agriculture majoring in Statistics. Faculty member at Department of Mathematics The University of Jember, Indonesia. afhadi.fmipa@unej.ac.id

²Professor of Biometrics Department of Statistics. Bogor University of Agriculture. Indonesia

Abstract. AMMI (Additive Main Effect Multiplicative Interaction) model for interactions in two-way table provide the major mean for studying stability and adaptability through genotype \times environment interaction (GEI), which modeled by full interaction model. Eligibility of AMMI model depends on that assumption of normally independent distributed error with a constant variance. Nowadays, AMMI models have been developed for any condition of MET data which is violence the normality, homogeneity assumption. We can mention in this class of modelling as M-AMMI for mixed AMMI models, G-AMMI for generalized AMMI models. The G-AMMI handles non-normality i.e categorical response variables using an algorithm of alternating regression. While in handling the non-homogeneity in mix-models sense, one may use a model called factor analytic multiplicative. The development of AMMI models is also to handle any outlier that might be found coincides with non-homogeneity condition of the data. In this paper, we will present methods of handling outlier in multiplicative model by robust approach of alternating regression algorithm.

Keyword: AMMI, G-AMMI, M-AMMI, Factor Analytic, Multiplicative Models, Alternating Regression, Robust approach.

1. Introduction

AMMI (Additive Main Effects and Multiplicative Interaction Analysis) was well described by Gauch (1988, 1992) and Gollob (1968). AMMI may be viewed as a procedure to separate pattern (the $G \times E$ interaction) from noise (mean error of treatment mean within trials). This is achieved by PCA, where the first axes (i.e. the axes with the largest eigenvalues) recover most of the pattern, whilst most of the noise ends up in later axes. The pattern can be viewed as the whole $G \times E$ effects weighed by an

estimate of the pattern-to-noise ratio associated with the respective effect. This pattern-to-noise ratio is a variance component ratio analogue to a repeatability or heritability coefficient (Piepho, 1994). Multiplicative models AMMI have been popularised in a fixed model context and a number of applications have been found (Gauch 1988, Gauch 1992, Crossa et al. 1990). AMMI analysis combines, in a model, additive components for main effects (treatments and environments) and multiplicative components for $G \times E$ effects.

AMMI model combines a univariate technique, ANOVA for the main effects and a multivariate technique PCA-principal component analysis, for $G \times E$ effects. Crossa (1990) suggests that the use of multivariate techniques permits a better use of information than the traditional regression methods. These models also give a visually pattern of the main interaction through biplot. The power of multiplicative AMMI model is placed on visualized interactions by using biplot.

The AMMI model represents observations into a systematic component that consists of main effect and interaction effect through multiplication of interactions components, apart from random errors component. Basically, the AMMI analysis combines both additive analysis of variance for the main effect of treatment and analysis of multiple main components uses bilinear modeling for the interaction effect, by using singular value decomposition (SVD) of its interaction matrix (Mattjik & Sumertajaya 2000, Mattjik 2005). Eligibility of AMMI model depends on that assumption of normally independent distributed error with a constant variance. Nowadays, AMMI models have been developed for any condition of MET data which is violences in i.e the normality and homogeneity assumption. We can mention in this class of modelling as M-AMMI for mixed AMMI models, G-AMMI for generalized AMMI models. The G-AMMI was handling non-normality i.e categorical response variables using an algorithm of alternating regression (Eeuwijk 1995), The G-AMMI with normal data distributed and identity link function is equal to AMMI models. Some applications of generalized alternating regression specifically on poisson distribution can be found at Hadi, AF et.al, 2010.

Handling the non-homogeneity in mix-models sense of AMMI, one may use a model called factor analytic multiplicative (Smith, et al. 2002). He also conclude that the factor analytic models is equivalent with AMMI mixed models. The development of AMMI models is also to handle any outlier that might be found coincides with non-homogeneity condition of the data. We know that SVD is venerable to the outliers, so that the construction of robust model is a promising subject to be investigated. In this paper, we will present of handling outlier in multiplicative model by robust approach of alternating regression algorithm.

2. Factor Analytic ANOVA (FANOVA) Models and its Relation to AMMI Models

A model concerning the evaluation of several treatments or genotypes in several environments is given by:

$$y_{ij} = \mu + g_i + e_j + ge_{ij} + \varepsilon_{ij}$$

where: μ , g , e , ge and ε are the fixed constant, genotype, environment, genotype x environment interaction and within environment error effects, respectively. The μ and e effects can be regarded as fixed and the others as random. In the context of MET data, the factor analysis approach can be used to provide a class of structures for the variance-covariance matrix of g_{ij} (G). The model is postulated in terms of the unobservable genotype effects in different environments:

$$ge_{ij} = \sum_{r=1}^k \lambda_{jr} f_{ir} + \delta_{ij}$$

where: ge_{ij} is interaction effect of genotype i in environment j ; λ_{jr} is loading for factor r in environment j ; f_{ir} is score for genotype i in factor r and δ_{ij} is error representing the lack of fit of the model.

The factor analytic mixed-model is presented according to Smith et al. (2002). Applied to g genotype effects on s environments (Resende & Thompson 2005), the factor analytic model postulates dependence on a set of random hypothetical factors $f_r^{(s \times 1)}$ ($r=1, 2, \dots, k < s$). In vector notation, the factor analytic model for these effects is

$$ge_s = (\lambda_1 \otimes I_g) f_1 + \dots + (\lambda_k \otimes I_g) f_k + \delta$$

where: $\lambda_r^{(s \times 1)}$: loadings or weights of the factors in environments; $\delta^{(gs \times 1)}$: vector of residuals or lack of fit for the model (also called vector of specific factors). In a compact way, the model is: $ge = \Lambda \otimes I f + \delta_{sg}$, where: $\Lambda^{(s \times k)} = [\lambda_1, \dots, \lambda_k]$; $f^{(gk \times 1)} = (f_1', \dots, f_k')$.

The joint distribution of f and δ is given by

$$\begin{pmatrix} f \\ \delta \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} I_k \otimes I_m & 0 \\ 0 & \Psi \otimes I_m \end{pmatrix} \right]$$

where: $\Psi = \text{diag}(\psi_1, \dots, \psi_p)$; ψ_i : specific variance for the i^{th} trial. The variance matrix for genotype effects on environments is given by $\text{var}[ge_s] = (\Lambda \otimes I_m) \text{var}[f] (\Lambda' \otimes I_m) + \text{var}(\delta) = (\Lambda \Lambda' + \Psi) \otimes I_m$. The model for genotype effects in each environment leads to a model for G in which:

$$\sigma_{g_{jj}} = \sum_{r=1}^k \lambda_{jr}^2 + \psi_j$$
 : genotype variance in environment j ;

$$\sigma_{g_{jj'}} = \sum_{r=1}^k \lambda_{jr} \lambda_{j'r}$$
 genotype covariance between environments j and j' ;

$$\rho_{g_{jj'}} = \frac{\sum_{r=1}^k \lambda_{jr} \lambda_{j'r}}{\sqrt{(\sum_{r=1}^k \lambda_{jr}^2 + \psi_j)(\sum_{r=1}^k \lambda_{j'r}^2 + \psi_{j'})}}$$
 : genotype correlation between environments j and j'

The equation for g_s has the form of a (random) regression on k environmental covariates $\lambda_1, \dots, \lambda_k$ in which all regressions pass through the origin. It may be more appropriate to allow a separate (non-zero) intercept for each genotype. This is equivalent to the model

Model AMMI Kini dan Yang Akan Datang

with genotype main effects, g , and a k -factor analytic model for $g \times e$ interaction. Then, the expression for g_s turns to

$$g_s = (1_s \otimes I_g)g + ge = (1_s \otimes I_g)g + (\Lambda \otimes I_g)f + \delta.$$

Vector g has mean zero and variance $\sigma_g^2 I$ or $\sigma_g^2 A$, where A is a genetic relationship matrix. The model can be written as

$$g_s = (\sigma_g 1_s \otimes I_g)f_0 + (\Lambda \otimes I_g)f + \delta = (\Lambda_g \otimes I_g)f_g + \delta$$

where:

$$\Lambda_g^{s(k+1)} = [\sigma_g 1_s \Lambda]; f_0 = g/\sigma_g; f'_g = (f'_0 f'_0')$$

Thus the model with genotype main effects and a k -factor analytic model for $g \times e$ interactions is a special case of a $(k+1)$ -factor analytic genotype effects in each environment, in which the loadings in the first set are constrained to be equal. The feature that distinguishes equations for g , from standard random multivariate regression problems is that both the covariates and the regression coefficients are unknown and therefore must be estimated from the data.

The AMMI model has become a popular method for analysing MET data. The model is a fixed-effects model with (additive) main effects for genotypes and environments and multiplicative terms for the interaction. The latter are obtained using a singular value decomposition (SVD) of the $G \times E$ interactions.

Let U_{ve} denote the $m \times p$ matrix of $G \times E$ interactions. In AMMI, U_{ve} is decomposed as $U_{ve} = \mathbf{A}\mathbf{L}\mathbf{B}^*$, where \mathbf{A} and \mathbf{B}^* are $m \times t$ and $p \times t$ matrices, such that $\mathbf{A}'\mathbf{A} = \mathbf{I}_t = \mathbf{B}^{*'}\mathbf{B}^*$, $\mathbf{L} = \text{diag}(l_1 \dots l_t)$, and t is the rank of U_{ve} .

Defining $\mathbf{B} = \mathbf{B}^*\mathbf{L}$, the decomposition can be written as

$$U_{ve} = \mathbf{A}\mathbf{B}' = \sum_{r=1}^t a_r b_r'$$

The columns of A (a_r of $m \times 1$) are called the genotype scores and the columns of B (b_r of $p \times 1$) are the environment loadings. As in factor analysis, the aim of the AMMI approach is to account for structure in the genetic effects using the minimum number, k , of multiplicative terms. Isolation of the first k terms in

$\text{var}[g_{e_s}] = (\Lambda \otimes I_m) \text{var}[f](\Lambda' \otimes I_m) + \text{var}(\delta)$ Then we have:

$$U_{ve} = \sum_{r=1}^k a_r b_r' + \sum_{r=k+1}^t a_r b_r' = A_1 B_1' + A_2 B_2'$$

Where A_1 and B_1 are $m \times k$ and $p \times k$ matrices, respectively. Thus, in the AMMI model, the $G \times E$ interactions are modeled as

$$\mathbf{u}_{ve} = (\mathbf{B}_1 \otimes \mathbf{I}_m)\mathbf{a} + \mathbf{e}_g$$

where $\underline{\mathbf{a}} = \text{vec}[A_1] = (a'_1 \dots a'_1)$ of $mk \times 1$ vector and $\underline{\mathbf{e}}_g$ of $mp \times 1$ are the residual $G \times E$ interactions that remain if not all t components of the SVD are used. The latter are assumed to be independent with constant variance. There is a clear connection between $\underline{\mathbf{u}}_{ve} = (\mathbf{B}_1 \otimes \mathbf{I}_m)\underline{\mathbf{a}} + \underline{\mathbf{e}}_g$ and the k factor-analytic model for the $G \times E$ interactions, namely:

$$\underline{\mathbf{u}}_{ve} = (\Lambda \otimes \mathbf{I}_m)\underline{\mathbf{f}} + \underline{\boldsymbol{\delta}}$$

There is a correspondence between the environment loadings for the two models (\mathbf{B}_1 and Λ) and the genotype scores ($\underline{\mathbf{a}}$ and $\underline{\mathbf{f}}$). Thus, the k factor-analytic model is a random-effects analogue of the AMMI model. The model is then a multiplicative model of environment and genotypes coefficients (known as loadings and factorial scores, respectively). Here, the FANOVA models is analogue to AMMI models.

In historical perspectives the FANOVA model is proposed by Gollob in 1968, combines aspects of analysis of variance and factor analysis. Among others, Gabriel (1978) considered models like FANOVA and estimated the unknown parameters using a least squares fit. AMMI then developed afterwards. Cornelius, et al. 1996 say that Gaugh & Zobel renamed the Gollob FANOVA model as AMMI model.

3. The RAR Estimator

As usual, the $n \times p$ data matrix Y contains the individuals (cases, objects) in the rows and the observed variables (characteristics) in the columns. The variables are already standardized to have zero location and unit spread. A factor score is denoted as f_{ij} . The i th score vector is given by $f_i = (f_{i1}, \dots, f_{ik})^T$, while the j th loading vector is $\lambda_j = (\lambda_{j1}, \dots, \lambda_{jk})^T$. Both the loading vectors and the score vectors are unknown. Denote by $\theta = (f_1^T, \dots, f_n^T, \lambda_1^T, \dots, \lambda_p^T)$ the vector of all scores and loadings, and let

$$\hat{y}_{ij}(\theta) = \sum_{l=1}^k f_{il}\lambda_{lj} = f_i^T \lambda_j = \lambda_j^T f_i$$

be the fitted value of y_{ij} . By choosing θ such that the fitted and the actual values of the data matrix are close together, we define estimates \hat{f}_i the score vectors and $\hat{\lambda}_j$ for the loading vectors. The fitted data matrix \hat{Y} can then be decomposed as $\hat{Y} = \hat{F}\hat{\Lambda}^T$ where the rows of \hat{F} are the estimated scores and the rows of $\hat{\Lambda}^T$ are the estimated loadings.

Observe that the rank of \hat{Y} is at most $k < p$, while the rank of Y is typically p . The least squares (LS) approach is to minimize the sum of squared residuals:

$$\hat{\theta}_{LS} = \underset{\theta}{\text{argmin}} \sum_{i=1}^n \cdot \sum_{j=1}^p (y_{ij} - \hat{y}_{ij}(\theta))^2$$

The resulting \hat{Y} can be seen as the “best” (in the least squares sense) approximation of the data matrix Y by a rank k matrix. The Eckart-Young theorem (Gower and Hand 1996) says that this best fit can be obtained by performing a singular value decomposition $Y = UDV^T$ of the data matrix. By replacing all singular values in D by zero except for the k largest ones, one obtains D_k and finally $\hat{Y} = UD_kV^T$. By taking $\hat{F} = \sqrt{n}U$ and $\hat{\Lambda} = VD_k/\sqrt{n}$ we obtain the so-called *Principal Component solution* to the FA problem

(Johnson and Wichern 1998). Moreover, the sample covariance matrix of the estimated score vectors equals $\hat{F}^T \hat{F} = I_k$ which is consistent with the assumption $Cov(F) = I_k$.

It is important to note that the estimates \hat{F} and $\hat{\Lambda}^T$ are only specified up to a linear transformation. Since $\hat{Y} = (\hat{F}^T T^T)(\hat{\Lambda}^T)^T$ for any non singular k by k matrix T , it follows that $\hat{F}^T T^T$ and $\hat{\Lambda}^T T^{-1}$ attain the same value for the objective $\hat{\theta}_{LS}$. However, the fitted values \hat{Y} are uniquely defined. Moreover, if we add the restriction that the estimated covariance matrix of the score vectors needs to be the identity matrix, then the estimates \hat{F} and $\hat{\Lambda}^T$ are specified up to an orthogonal transformation, making the matrix $\hat{\Lambda} \hat{\Lambda}^T$ uniquely defined.

Since the LS criterion gives too much weight to large residuals, a first idea is to use the L_1 criterion (or Least Absolute Deviations criterion) instead, which is known to give a very robust additive fit to two-way tables (Terbeck & Davies 1998). This yields the L_1 estimator below

$$\hat{\theta}_{L_1} = \operatorname{argmin}_{\theta} \sum_{i=1}^n \sum_{j=1}^p |y_{ij} - \hat{y}_{ij}(\theta)|.$$

For the optimal \hat{F} and $\hat{\Lambda}^T$, it must hold that \hat{f}_i minimizes $\sum_{j=1}^p |y_{ij} - f_i^T \hat{\lambda}_j|$ and $\hat{\lambda}_j$ minimizes $\sum_{i=1}^n |y_{ij} - \hat{f}_i^T \lambda_j|$. Therefore, instead of minimizing both sums in L_1 criterion at the same time, one fixes an index j and scores f_i and selects the λ_j to minimize $\sum_{i=1}^n |y_{ij} - f_i^T \lambda_j|$

The above problem is now linear instead of bilinear and can easily be solved with a Least Absolute Deviations regression algorithm. One sees immediately that minimizing $\sum_{i=1}^n |y_{ij} - f_i^T \lambda_j|$ consecutively for $j = 1, \dots, p$ corresponds to minimizing L_1 criterion for fixed scores. Analogously, for fixed loadings λ_j , finding the f_i minimizing $\sum_{j=1}^p |y_{ij} - f_i^T \lambda_j|$ (for each $i = 1, \dots, n$ in turn) corresponds to minimizing L_1 criterion when the loadings are given. Alternating $\sum_{i=1}^n |y_{ij} - \hat{f}_i^T \lambda_j|$ and $\sum_{j=1}^p |y_{ij} - \hat{f}_i^T \lambda_j|$ leads to an iterative scheme of alternating regressions. Note that the value of the criterion in L_1 decreases at each step.

Similar algorithms, but based on alternating classical least squares regressions and generalized linear models, see de Falguerolles and Francis 1992, Gabriel 1998 for generalized bilinear models.

Unfortunately, L_1 regression is sensitive to leverage points. If outlying score or loading vectors are present, the L_1 regressions can be heavily influenced by them. By downweighting these leverage points we obtain a weighted L_1 regression, resulting in the

$$\operatorname{argmin}_{\theta} \sum_{i=1}^n \cdot \sum_{j=1}^p w_i(\theta) v_j(\theta) |y_{ij} - \hat{y}_{ij}(\theta)| \quad \hat{\theta}_{RAR} =$$

One single objective function estimates \hat{F} and $\hat{\Lambda}^T$ simultaneously from the rows and columns of Y . The result of $\hat{\theta}_{RAR}$ used Robust Alternating Regressions to compute it. The estimator will not be misled by outlying observations.

The row weights, $w_i(\theta)$ are defined by $w_i(\theta) = \min(1, \frac{\chi_{k,0.95}^2}{RD_i^2})$ for $i = 1, \dots, n$ where $\chi_{k,0.95}^2$ is the upper 5% critical value of a chi-squared distribution with k degrees of freedom, and

$$RD_i = \sqrt{(f_i - T(F))^T C(F)^{-1} (f_i - T(F))}$$

for $i = 1, \dots, n$ are robust distances (Rousseeuw and van Zomeren 1990) computed from the collection of score vectors $F = \{f_i | 1 \leq i \leq n\}$ k -dimensional space. Such weights were used by Simpson et al (1992) and yielded stable results. The robust multivariate location and scatter estimators T and C are taken as the location and scatter part of the MVE estimator. The MVE estimator was chosen here since it performs well as an outlier identifier (see Becker and Gather 2001). Analogously, the set of column weights v_j is defined using the loading vectors. Note that, since the true loadings and scores are unobserved, w_i and v_j depend on the unknown parameter vector θ .

4. The RAR Algorithm

The RAR estimator can be approximated by an alternating algorithm, as outlined below (C. Croux, et al. 2005).

Step 0: To obtain invariance with respect to a change of measurement units, the data are first scaled in a robust way:

$$y_{ij} \leftarrow \frac{y_{ij} - med_i(y_{ij})}{MAD_i(y_{ij})}$$

where MAD stands for the Median Absolute Deviation. Note that orthogonal or affine equivariance properties are not necessary in a factor model. This initial standardization corresponds with a correlation matrix based FA.

Step 1: Starting values. First, a robust principal component analysis (PCA) procedure is performed. The resulting scores are then taken as starting values $\hat{f}_i^{(0)}$ for the factor scores. We use the projection pursuit (PP) based estimator of Li and Chen (1985), implemented as in Croux and Ruiz-Gazen (1996). This PP-based method is fast to compute, can deal with $p > n$, and is highly robust. Moreover, this approach allows one to compute just the first k principal components (the only ones that are needed here), which reduces the computation time even further. Using classical PCA in this first stage would slow down the convergence considerably, and could lead to a nonrobust FA when there are many outliers. Alternatively, one could take several random starting values, which could help to check for a local versus global optimum. But the latter approach will increase computation time significantly. In any case, experiments have shown that the choice of the starting values is not too crucial for finding a good approximation.

Step 2: The iteration process. Now suppose that the iteration process has reached step t ($t \geq 1$) of the algorithm, and the $\hat{f}_i^{(t-1)}$ are available.

* First compute the row weights $w_i^{(t)}$, which downweight outliers in the set of estimated score vectors

$$\{\hat{f}_i^{(t-1)} | 1 \leq i \leq n\} \text{ in } \mathbb{R}^k.$$

Then compute

$$\hat{\lambda}_j^{(t)} = \operatorname{argmin}_{\lambda \in \mathbb{R}^k} \sum_{i=1}^n w_i^{(t)} |y_{ij} - \lambda^T \hat{f}_i^{(t-1)}|$$

for $j = 1, \dots, p$: In this part of the procedure, one needs to perform an L_1 fit p times (and this will be the case at every iteration step). Note that the loadings are estimated one at a time, which turned out to be more convenient for the implementation of the algorithm. Fortunately, very efficient algorithms for L_1 regression exist (Bloomfield and Steiger 1983), so this takes little time. Note that the weights $w_i^{(t)}$, only need to be computed once every iteration step. They require computation of a robust scatter estimator in the factor space, which is usually of a low dimension k .

* We analogously compute column weights $v_j^{(t)}$, which downweight outliers in the set of estimated loading vectors $\{\hat{\lambda}_i^{(t)} | 1 \leq j \leq p\}$ in \mathbb{R}^k . Then compute

$$\hat{f}_i^{(t)} = \operatorname{argmin}_{f \in \mathbb{R}^k} \sum_{j=1}^p v_j^{(t)} |y_{ij} - f^T \hat{\lambda}_j^{(t)}|$$

for $i = 1, \dots, n$.

* The values of the objective function of $\hat{\theta}_{RAR}$ computed for the estimates obtained in step $t-1$ and step t are compared. If there is no essential difference in the objective function, the iterative process is stopped and we set $\hat{f}_i = \hat{f}_i^{(t)}$ for $1 \leq i \leq n$ and $\hat{\lambda}_j = \hat{\lambda}_j^{(t)}$ for $1 \leq j \leq p$. If not, Step 2 is repeated.

Step 3: Orthogonalization. This last step is optional and will not alter the fitted values $\hat{Y} = \hat{F}\hat{\Lambda}^T$. We compute a robust estimator $\hat{\Sigma}_f$ of the covariance matrix of the estimated scores $\{\hat{f}_i | 1 \leq i \leq n\}$. Since the scores only have dimension k , where k is small, the matrix $\hat{\Sigma}_f$ can be computed quickly. We compute \hat{f}_i by the reweighted MCD estimator with 25% breakdown value, using the FAST-MCD algorithm of Rousseeuw and van Driessen (1999). The breakdown value 25% for the MCD has been chosen since this combines robustness with efficiency (see e.g. Croux and Haesbroeck 1999).

Afterwards we set

$$\hat{F} \leftarrow \hat{F}\hat{\Sigma}_f^{-1/2} \text{ and } \hat{\Lambda} \leftarrow \hat{\Lambda}\hat{\Sigma}_f^{1/2}$$

The effect of the above transformation is that the robust covariance matrix of the estimated scores is now an identity matrix, which mimics the model condition $Cov(\hat{F}) = I_k$. Another effect is that the biplot representation of the n cases (see Step 4) will show no correlation structure, as is common practice in the biplot literature (Gower & Hand 1996).

Step 4: Residuals, uniquenesses, biplot. The residuals are obtained as $\hat{\varepsilon}_{ij} = y_{ij} - \hat{y}_{ij} = y_{ij} - \hat{f}_i^T \hat{\lambda}_j$ and can be plotted versus (i, j) in the horizontal plane. This residual plot is very useful for detecting outliers. From the residuals the uniquenesses can

be estimated as $\hat{\psi}_j = (MAD_j(\hat{\varepsilon}_{ij}))^2$. In the common case $k = 2$ one can represent the individuals by $(\hat{f}_{i1}, \hat{f}_{i2})$ and the variables by $(\hat{\lambda}_{j1}, \hat{\lambda}_{j2})$ in the same 2D plot, called the biplot.

It also allows to perform alternating regression using other regression estimators, like M-estimators or the highly robust Least Trimmed Squares (LTS) and Least Median of Squares (LMS) estimators. It is even possible to execute the algorithm with the nonrobust Least Squares regression estimator, yielding the same result as the classical approach of Gabriel (1978) based on the singular value decomposition. Alternating regression using the LMS algorithm was already considered by Ukkelberg and Borgen (1993). However, using the LMS yields a very time consuming algorithm. In our experience, the RAR estimator gave the most satisfying factor analysis method with respect to computation time, robustness, and stable convergence of the algorithm. Although no proof of convergence exists, many simulations and examples have shown its good numerical and statistical performance.

The RAR procedure required the choice of several auxiliary robust estimators and a weighting function. Most of these choices are standard, and simulations for other robust choices led to essentially identical results.

5. Developing Robust AMMI by applying RAR to the FANOVA Models

The standard model for a two-way table is the ANOVA model $y_{ij} = \mu + g_i + e_j + \delta_{ij}$ where μ is called the overall mean, g_i represents genotype effect (the row) and e_j environment (the column) effect. In a classical setup, the row and column effects are assumed to have zero mean. The terms δ_{ij} can either be seen as residuals or as interaction terms between rows and columns. That expression is called an additive model. It is however quite possible that the interaction terms δ_{ij} still contain some structure that can be described by a factor model $\delta_{ij} = \sum_{r=1}^k \lambda_{jr} f_{ir} + e_{ij}$ yielding the overall model

$$y_{ij} = \mu + g_i + e_j + \sum_{r=1}^k \lambda_{jr} f_{ir} + \varepsilon_{ij}$$

A first idea would be to proceed sequentially by estimating the additive model first, and afterwards performing a factor analysis on the residuals. But better fits can be obtained by estimating all parameters jointly. For the least squares fit there is no difference between the simultaneous and the sequential approach, but this is no longer true for the robust fits. Therefore we will estimate additive and multiplicative terms simultaneously.

The RAR estimator for the FANOVA model can be defined as in Section 2. Denote θ as the vector collecting the scores, loadings, row and column effects and the overall effect μ . In order to estimate the $(k + 1)(n + p) + 1$ unknown elements of θ from the np available data, we can use the RAR estimator defined as in section 3:

$$\hat{\theta}_{RAR} = \underset{\theta}{\operatorname{argmin}} \sum_{j=1}^p \sum_{i=1}^n v_j(\theta) w_i(\theta) |y_{ij} - \hat{y}_{ij}(\theta)|$$

The weights w_i and v_j are defined as in section 3 and are downweighting outlying scores and loadings in the k -dimensional spaces of scores and loadings. To uniquely identify the parameters in y_{ij} , the function $\hat{\theta}_{RAR}$ will be minimized under the constraints:

$$\text{med}_i(g_i) = \text{med}_j(e_j) = 0 \text{ and}$$
$$\text{med}_i(f_{il}) = \text{med}_j(\lambda_{jl}) = 0 \text{ for } l = 1, \dots, k$$

Those constraints are consistent with robust approach. The algorithm to compute the RAR estimator in FANOVA models is based on alternating regressions, and is almost identical to the iterative scheme outlined in Section 4. The difference is only, it is no longer working with regression through the origin, intercepts need to be estimated.

6. Illustrative Example of Fitting Robust AMMI by Robust alternating FANOVA Models

The data we will use is national research conducted by Indonesian National Consortium of Rice (INCR) in cooperation with Indonesian Centre for Rice Research in 2008. The aim of this research was to evaluate the phenotypic performance of rice from the latest generation in the different environment. There are 11 cultivars evaluated at 20 environments. There are 3 cultivars from BATAN, 4 cultivars from ICRR, 2 cultivars from Biogen, and 2 cultivars from IPB, with 3 varieties to compare (Ciherang, Inpari1, Cimelati). It used 3 plots of replication for each cultivar in each environment, sized in 4 m x 5 m. We will use two way table of genotype \times environment interaction, the cell was median of 3 plots of replication.

Result. There were no super outlier in the data, only one outlier identified in the row effect (Figure 2). So it can be predicted that the result will be slightly different with the least square criterion. From figure 1, it can be said that L4 have relatively largest variance.

The genotype which coordinates near the centre point (0,0) is clustered as a stable genotype, so G3, G8, G4, G12, and G14 are relatively more stable than other genotypes. When there were no super outlier then the robust methods give the slightly different result from the least square ones. Figure 2 shows that only one observation that suspected to be an outlier in column effect. And its downweighted already by shown in Tabel 1, that was in the L4. The residual plot (Figure 3) shows that its random distributed by row and column way.

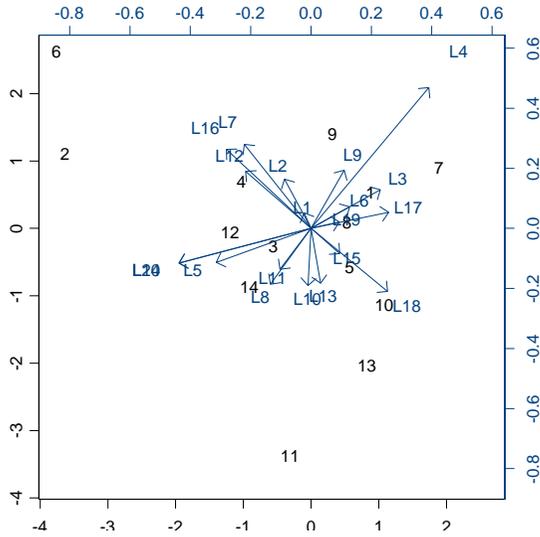


Figure 1. Biplot of $G \times E$ interaction of Robust FANOVA Models (factor 1 49.43%, cumulative 2 factor 53.76%)

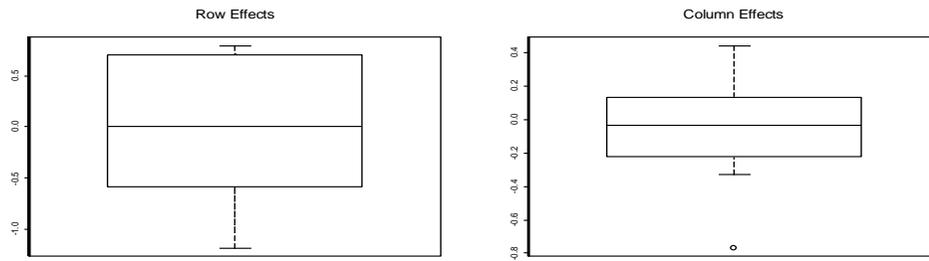


Figure 2. The Boxplot of Row (Environment) and column (Genotype) effect in INCR 2008 data

Table 1. Row and column downweight to the outlier(s)

Row Weight:				Column Weight:					
G1	1	G8	1	L1	1	L8	1	L15	1
G2	1	G9	1	L2	1	L9	1	L16	1
G3	1	G10	1	L3	1	L10	1	L17	1
G4	1	G11	1	L4	0.698925	L11	1	L18	1
G5	1	G12	1	L5	1	L12	1	L19	1
G6	1	G13	1	L6	1	L13	1	L20	1
G7	1	G14	1	L7	1	L14	1		

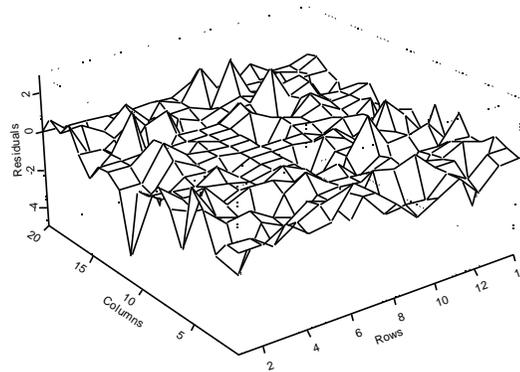


Figure 3. Residual plots of robust FANOVA Models

7. Summary and Conclusions

FANOVA model of Gollob is combines aspects of analysis of variance and factor analysis and estimated the unknown parameters using a least squares fit. AMMI then developed afterwards. Gaugh & Zobel renamed the Gollob FANOVA model as AMMI model. AMMI model combines a univariate technique (ANOVA) for the main effects and a multivariate technique (PCA and SVD) for $G \times E$ effects. While in handling the non-homogeneity in mix-models sense, one may use a model called factor analytic multiplicative. Factor analitic models is equivalent wth AMMI mixed model. SVD is venerable to the outliers, so that the construction of robust model is a promising subject to be investigated. In this paper, we will present of handling outlier in multiplicative model by robust approach of alternating regression algorithm.

Basically the idea of constructing robust FANOVA model was to proceed sequentially by estimating the additive effect by robust model first, and afterwards performing a robust factor analysis on the residuals. However, this is no longer true for the robust fits. Therefore we will estimate additive and multiplicative terms simultaneously by applying Robust Alternating Regression in FANOVA model. RAR FANOVA model is are downweighting outlying scores and loadings in the k -dimensional spaces of scores and loadings, and uniquely identify the parameters in y_{ij} , the function $\hat{\theta}_{RAR}$ will be minimized under the constraints that are consistent with robust approach of the median of parameters. When there were no sufer outlier then the robust methods give the slightly different result from the least square ones.

8. Reference

Becker C & Gather U. 2001. The largest nonidentifiable outlier: A comparison of multivariate simultaneous outlier identification rules. *Computational Statistics and Data Analysis*. **36**:119-127.

- Bloomfield P & Steiger WL. 1983. *Least Absolute Deviations: Theory, Applications, and Algorithms*, Boston, Mass, Birkh user.
- Cornelius PL, Crossa J, & Seyedsadr MS. 1996. Statistical Test and Estimators of Multiplicative Models for GEI. In *Genotype by Environment Interaction*. Manjit S. Kang & Gauch HG (eds). CRC Press. (pp. 200) <http://books.google.co.id/books?id=NPqiC-mvRTUC&lpg=PP1&pg=PA199> Visit at 10 July 2009.
- Croux C & Haesbroeck G. 1999. Influence function and efficiency of the minimum covariance determinant scatter matrix estimator. *Journal of Multivariate Analysis*. **71**:161-190.
- Croux C & Haesbroeck G. 2000. Principal component analysis based on robust estimators of the covariance or correlation matrix: influence functions and efficiencies, *Biometrika*. **87**:603-618.
- Croux C & Ruiz-Gazen A. 1996. A fast algorithm for robust principal components based on projection pursuit. *COMPSTAT 1996, Proceedings in Computational Statistics* (ed. A. Prat). Heidelberg, Physica-Verlag. pp. 211-216.
- Crossa J. 1990. Statistical Analysis of Multilocation Trials. *Adv. Agron*. **44**: 55-85.
- Croux C, Filzmoser P, Pison G & Rousseeuw PJ. 2003. Fitting Multiplicative Models by Robust Alternating Regressions.
- de Falguerolles A & Francis B. 1992. Algorithmic approaches for fitting bilinear models. *COMPSTAT 1992, Proceedings in Computational Statistics, Vol. 1* (eds. Y. Dodge and J. Whittaker), Heidelberg, Physica-Verlag. pp. 77-82.
- Denis J B & Gower J C. 1996. Asymptotic confidence regions for biadditive models: Interpreting genotype-environment interactions. *Applied Statistics*. **45**:479-493.
- Gabriel K R. 1978. Least squares approximation of matrices by additive and multiplicative models, *Journal of the Royal Statistical Society B*, **40(2)**:186-196.
- Gabriel K R. 1998. Generalized bilinear regression, *Biometrika*, **85**:689-700.
- Gabriel K R & Zamir S. 1979. Lower rank approximation of matrices by least squares with any choice of weights, *Technometrics*, **21**:489-498.
- Gauch H G. 1988. Model selection and validation for yield trial with interaction. *Biometrics*. **44**:705-716.
- Gollob H F. 1968. A statistical model which combines features of factor analytic and analysis of variance techniques. *Psychometrika*. **33**:73-116.
- Gower J & Hand D. 1996. *Biplots*, New York, Chapman & Hall.
- Hadi A F, Mattjik A A, & Sumertajaya IM. 2010. Generalized AMMI Models For Assessing The Endurance of Soybean to Leaf Pest. *Jurnal Ilmu Dasar*. **11(2)**: 151-159.
- Johnson R A & Wichern D W. 1998. *Applied Multivariate Statistical Analysis*. 4th ed. Prentice Hall. New Jersey.
- Li G & Chen Z. 1985. Projection-pursuit approach to robust dispersion matrices and principal components: Primary theory and Monte Carlo. *Journal of the American Statistical Association*. **80**:759-766.

Model AMMI Kini dan Yang Akan Datang

- Mattjik, A A & Sumertajaya I M. 2000. *Perancangan Percobaan dengan Aplikasi SAS dan Minitab*. IPB Press, Bogor.
- Mattjik, A A. 2005. *Interaksi Genotipe dan Lingkungan dalam Penyediaan Sumberdaya Unggul*. [Naskah Orasi Ilmiah Guru Besar Biometrika. Tidak Dipublikasikan. Institut Pertanian Bogor. Bogor].
- Resende, M D P & Thompson R. 2004. Factor Analytic Multiplicative Mixed Model in The Analysis of Multiple Experiments. *Rev. Mat. Estat.* 22(2): 31-52.
- Rousseeuw P J & Van Driessen K. 1999. A fast algorithm for the minimum covariance determinant estimator, *Technometrics*, **41**:212-223.
- Smith A, B. Cullis D, Lockett G, Hollamby & Thompson R. 2002. Exploring genotype-environment data using random effects AMMI models with adjustments for spatial field trend: *Part 1-2: Theory-Applications*. In. Kang, M.S. (Ed). *Quantitative Genetics, Genomics and Plant Breeding*. CAB International, UK
- Terbeck W & Davies P. 1998. Interactions and outliers in the two-way analysis of variance, *The Annals of Statistics*. **26**:1279-1305.
- Ukkelberg °A & Borgen O. 1993. Outlier detection by robust alternating regressions. *Analytica Chimica Acta*. **277**:489-494.
- Van Eeuwijk F A 1995. Multiplicative Interaction in Generalized Linear Models. *Biometrics*. **51**:1017-1032.

9. Acknowledgment

1. This article is part of reseach that funded by the Ministry of National Education, The Republic of Indonesia, Directorate Generale of Higher Education at the HIBAH PENELITIAN PASCASARJANA, 2008-2010
2. Dr. Ir. Aan Darajat, Senior Researcher at Indonesian Center for Rice Research, for the data and the discussions.
3. Prof. Peter Filzmoser. Department of Statistics and Probability Theory. Vienna University of Technology for the S-PLUS scripts and the articles.

9

KETAKNORMALAN DATA PADA MODEL AMMI

Analisis stabilitas dapat digunakan untuk mengidentifikasi ketahanan terhadap hama dan penyakit. Jika ada interaksi antara kultivar dan patogen, maka perlu untuk mengidentifikasi suatu kultivar yang memiliki resistensi umum dan resistensi spesifik (Mattjik, 2005). Namun untuk kajian ini (ketahanan terhadap penyakit dan kejadian serangan hama pada tanaman misalnya), kelayakan model AMMI dengan galat yang Normal dan ragam konstan tidak selalu dapat dipenuhi. Jika matriks data bebas, bersebaran Normal dengan ragam konstan, penduga kemungkinan maksimum tereduksi menjadi SVD, namun tidak demikian bila tidak menyebar Normal.

Pencatatan data serangan hama/penyakit dalam bentuk cacahan (*counting*), atau data dalam persentase merupakan contoh fenomena ini. Sebagian peneliti melakukan pencatatan menggunakan cacahan kemudian menganalisisnya sebagai intensitas serangan dalam persentase. Dengan data persentase ini kemudian ia melakukan transformasi sehingga asumsi-asumsi yang diperlukan (kenormalan dan kehomogenan ragam) terpenuhi. Selama transformasi berhasil mengatasinya cukuplah bagi kita untuk memilih cara ini. Sayangnya menemukan transformasi yang sesuai tidaklah mudah.

Sementara itu, pada pemodelan aditif telah dikenal luas apa yang disebut dengan *Generalized Linear Models* (GLM) atau Model Linier Terampat (MLT) sebuah kelas pemodelan yang menangani data-data berdistribusi bukan Normal. Model multiplikatif (bilinear) menjembatani kesenjangan antara model pengaruh utama (pada ANOVA ataupun GLM) dan model interaksi lengkap dengan sebuah parameter interaksi untuk setiap sel dalam tabel dua arah. Model ini pun memberikan visualisasi corak utama interaksi melalui biplot. Karenanya pengembangan teori model linier terampat dengan mengakomodasi komponen multiplikatif untuk interaksi sangat diperlukan.

Kekuatan eksplorasi model multiplikatif AMMI terletak pada visualisasi interaksi melalui biplot. Van Eeuwijk, 1995, memperkenalkan model multiplikatif dalam konteks model linier terampat sebagai perluasan dari model AMMI yang disebut dengan Generalized AMMI atau disingkat GAMMI. Pada pemodelan GAMMI, visualisasi interaksi ini masih dimungkinkan.

Permasalahan yang muncul kemudian adalah bagaimana pengepasan model GAMMI ini dilakukan? Bagaimana dengan penggunaan transformasi sebagai upaya pemenuhan asumsi kenormalan pada model AMMI? Pendekatan manakah yang lebih menguntungkan? Penelitian ini ingin membandingkan kedua pendekatan tersebut, disertai penerapannya pada penelitian pemuliaan tanaman.

9.1 Pendekatan Transformasi Kenormalan

Analisis AMMI adalah suatu teknik analisis data percobaan dua faktor perlakuan dengan pengaruh utama perlakuan bersifat aditif sedangkan pengaruh interaksi dimodelkan dengan model bilinear. Model AMMI merepresentasikan observasi ke dalam komponen sistematis yang terdiri dari pengaruh utama (main effect) dan pengaruh interaksi melalui suku-suku multiplikatif (multiplicative interactions), di samping komponen acak sisaan atau galat. Komponen acak pada model ini

diasumsikan menyebar Normal dengan ragam konstan. Kelayakan model AMMI dengan galat yang Normal dan ragam konstan ada kalanya tidak terpenuhi. Transformasi data pengamatan mungkin menjadi salah satu teknik untuk mengatasi masalah ketidaknormalan ini.

Data yang berdistribusi bukan Normal cenderung tidak homogen, dan bila dimodelkan dengan AMMI ketakhomogenan ragam dapat berakibat buruk, sedangkan skala dugaannya mungkin juga tidak memuaskan. Kedua fenomena ini bisa jadi membutuhkan dimasukkannya suku interaksi tambahan (Van Eeuwijk, 1995). Kadangkala ada alasan kuat untuk tetap memodelkan data pada skala pengamatan. Kehomogenan ragam dapat diatasi dengan menambahkan satu atau lebih suku multiplikatif interaksi. Ketika tidak ada alasan untuk memaksa pemodelan tetap pada skala pengamatan, maka transformasi terhadap peubah respon dapat dilakukan untuk mengurangi masalah ini. Model linier atau bilinear dikenakan pada data yang telah ditransformasi, dan sifat sebaran sisaan diasumsikan memenuhi sebaran Normal. Transformasi pada peubah respon ditengarai merupakan upaya perbaikan atas kedua hal tersebut. Model linier atau bilinear dapat dikenakan pada data yang telah ditransformasi, dan sifat sebaran sisaan diasumsikan memenuhi sebaran Normal.

Transformasi data pada hakekatnya adalah suatu usaha untuk mengubah data dari suatu skala ke skala yang lain. Model linier yang klasik (analisis ragam atau regresi) telah dikembangkan berdasarkan pada beberapa asumsi pokok yaitu keaditifan (model pengaruh utama), ragam perlakuan yang homogen (keragaman data bersifat bebas dari rataan dan banyaknya ulangan), dan kenormalan data.

Asumsi pertama berkaitan dengan struktur data yang pada akhirnya menyangkut penafsiran data, asumsi kedua berperan dalam menyederhanakan metode pendugaan parameter. Sedangkan yang terakhir sangat erat kaitannya dengan pengujian hipotesis. Metode pengujian hipotesis yang telah berkembang sangat lanjut adalah yang didasarkan pada kenormalan data, oleh karena itu patokan-patokannya dapat dengan mudah diperoleh dalam tabel-tabel sebaran statistik, seperti tabel t, F atau Khi-kuadrat (Aunuddin, 2005).

Dalam hal ini, transformasi bertujuan untuk mengatasi tiga masalah utama yaitu keheterogenan ragam, ketaknormalan galat, dan ketakaditifan/ketaklinieran pengaruh sistematis. Diakui bahwa bagaimanapun, tidak mudah mengatasi ketiga hal tersebut dengan satu langkah tunggal transformasi. Transformasi tunggal biasanya manjur untuk mengatasi satu masalah tertentu tetapi tidak ketiganya. Keberhasilan transformasi untuk memperoleh kesederhanaan model (aditifitas/linieritas) mungkin mengakibatkan ketaknormalan dan

ketakhomogenan ragam bila sebelumnya dua asumsi ini terpenuhi. Ada kalanya transformasi yang dilakukan untuk memperoleh ragam yang statbil membawa kita pada ketaknormalan (Rawling, J.O. et al., 1998) .

Beruntunglah, bahwa transformasi untuk memperoleh kehomogenan ragam dan ketaknormalan mempunyai kecenderungan diperoleh secara bersamaan (hand-in-hand), sehingga tidak jarang kedua asumsi dapat terpenuhi oleh suatu transformasi yang tepat (Bartlett, 1947 diacu dalam Rawling, J.O. et al., 1998)

Transformasi untuk kehomogenan ragam seringkali juga memenuhi kenormalan. Transformasi logit, arcsinus, dan probit yang digunakan untuk menstabilkan ragam dan menyederhanakan model juga membuat distribusi mendekati kenormalan. Transformasi tersebut umumnya menarik (*stretching*) ekor distribusi untuk memberikan bentuk distribusi yang mendekati bentuk genta.

Demikian halnya dengan transformasi keluarga pangkat juga berguna untuk membuat distribusi menjadi semakin simetrik (mengurangi kemenjuluran). Harapannya adalah diperoleh distribusi data yang semakin mendekati Normal. Kriteria yang berbeda untuk menentukan transformasi apa yang akan digunakan tidak harus menuju pada pilihan yang sama, tetapi sering terjadi transformasi yang optimum untuk suatu masalah juga memperbaiki masalah yang lain. Pada keluarga transformasi ini telah dikenal luas suatu metode perhitungan untuk menentukan transformasi optimum, yaitu transformasi Box-Cox. Hadi, *et al.* 2007 melakukannya pendekatan transformasi Box-Cox ini.

Seperti dilaporkan Hadi, *et al* 2007, transformasi Box-Cox mampu mengatasi ketaknormalan data. Dengan transformasi Box-Cox dapat dilakukan pemodelan interaksi menggunakan model AMMI secara sah pada data ternormalkan. Namun transformasi kenormalan dilakukan semata-mata untuk tetap bertahan pada model dengan metode pendugaannya yang telah mapan secara teori sehingga pengujian hipotesis dan interpretasinya pun tidak banyak perdebatan. Model AMMI dibangun dengan landasan teori pemodelan yang mapan, teknik komputasi yang sederhana, dan telah secara luas digunakan. Transformasi data dilakukan untuk semata-mata memperoleh asumsi kenormalan. Analisis AMMI kemudian dilakukan pada data hasil transformasi ini. Kita seolah menutup mata terhadap makna apa yang diberikan oleh transformasi pada interpretasi model AMMI.

9.2 Model AMMI Terampat

Arti penting pemodelan statistika adalah menyediakan interpretasi atas fenomena yang dipelajari, dan menyatakannya dengan bahasa yang sesuai dengan bidang aplikasi. Transformasi dapat dihindari manakala kehomogenan ragam dapat dimodelkan oleh suku-suku multiplikatif pengaruh interaksi pada struktur sistematis model. Bagaimanapun, untuk data bukan Normal yang dimodelkan pada skala observasi, interaksi multiplikatif kemungkinan besar merefleksikan dua hal, kehomogenan ragam dan interaksi multiplikatif yang sebenarnya. Tidak ada jaminan penuh bahwa transformasi data pada skala pengamatan dapat memisahkan kedua hal di atas.

Transformasi, dalam kasus analisis regresi ataupun analisis ragam, bertujuan untuk memperoleh kehomogenan ragam, mendekati kenormalan galat, dan keaditifan pengaruh sistematis. Tidaklah mudah mendapatkan sebuah transformasi yang memenuhi semua kebutuhan itu. Sebagai contoh, untuk data cacahan yang berdistribusi Poisson dan pengaruh sistematisnya multiplikatif, transformasi akar akan berhasil memperoleh ragam yang konstan, transformasi pangkat dengan pangkat dua per tiga akan menghasilkan distribusi yang mendekati simetrik atau Normal, sedangkan transformasi logaritma menghasilkan aditifitas pengaruh sistematis. Jadi, setelah transformasi pun, suku multiplikatif kemungkinan (masih) mencerminkan campuran keheterogenan ragam dan pengaruh multiplikatif.

Dalam suatu percobaan, respon yang diamati terkadang berupa data kategorik. Hal ini mengakibatkan pendekatan model AMMI menjadi tidak relevan sehingga perlu dilakukan analisis dengan menggunakan pendekatan lain.

Pada pemodelan aditif telah dikenal luas apa yang disebut dengan Generalized Linear Models (GLM) sebuah kelas pemodelan yang menangani data-data non-normal. Model ini tidak bergantung pada karakteristik atau asumsi distribusi data (kenormalan dan kehomogenan ragam), tetapi bergantung hanya pada sifat fungsi hubung (*link function*) yang menghubungkan parameter rata-rata dan prediktor linear (*linear predictor*) dari model sebaran peluang yang digunakan. Namun model ini merupakan model aditif, sampai dengan tahun 1995, belum mengadopsi model interaksi lengkap seperti pada AMMI.

Eeuwijk, 1995 dan Falguolues, 1996 mengusulkan algoritma regresi bolak-balik untuk masing-masing faktor interaksi. Namun seperti dilaporkan Hadi *et al* 2008 algoritma ini diperkirakan akan menemui kompleksitas karena didalamnya akan terlibat sedikitnya 3 kekonvergenan. Di samping itu, algoritma ini menemui masalah utama ketakkonstanan ragam.

Model multiplikatif (bilinear) menjembatani kesenjangan antara model pengaruh utama (pada ANOVA atau GLM) dan model interaksi lengkap dengan satu parameter interaksi untuk tiap-tiap sel dalam tabel dua arah. Model ini pun memberikan visualisasi pola utama interaksi melalui Biplot. Karenanya pengembangan teori GLM dengan mengakomodasi komponen multiplikatif untuk interaksi sangat diperlukan. Introduksi model multiplikatif untuk kelas pemodelan GLM, sebagai generalisasi dari AMMI, disebut sebagai Generalized Linear Additive Main-Effect & Multiplicative Interactions Model (Model GAMMI).

Kekuatan model multiplikatif AMMI terletak pada visualisasi interaksi melalui Biplot. Pada model GAMMI, visualisasi interaksi ini masih dimungkinkan. Namun masih harus diinvestigasi karena sangat tergantung pada distribusi data dan link-function yang digunakan. Hadi, AF 2008 melaporkan bahwa *link-function* ini mempengaruhi interpretasi pada perbandingan peluang (odds) dan rasio odds yang dibutuhkan pada interpretasi. Karenanya pengembangan teori MLT dengan mengakomodasi komponen multiplikatif untuk interaksi sangat diperlukan. Sejauh ini beberapa penggunaan model AMMI terampat pada distribusi poisson dan binomial telah dilakukan sebagaimana oleh Hadi *et al* 2007, Hadi *et al* 2008.

9.3 GENERALIZED AMMI MODELS FOR ASSESSING THE ENDURANCE OF SOYBEAN TO LEAF PEST (Alfian F. Hadi, A. A. Mattjik & I M Sumertajaya, Jurnal Ilmu Dasar Terakreditasi. Vol 11. NO. 2. Juli 2010. P 151-159)

Generalized Ammi Models For Assessing The Endurance Of Soybean To Leaf Pest

Alfian F. Hadi¹, A. A. Mattjik², and I M Sumertajaya²

¹) *PhD Student Majoring in Statistics, Graduated School of Bogor Agricultural University
Faculty Member of Department of Mathematics, The University of Jember, email:*

afhadi.fmipa@unej.ac.id or afhadi@gmail.com

²) *Department of Statistics, Bogor Agricultural University*

Abstract

AMMI (Additive Main Effect Multiplicative Interaction) model for interactions in two-way table provide the major mean for studying stability and adaptability through genotype \times environment interaction (GEI), which modeled by full interaction model. Eligibility of AMMI (Additive Main Effect Multiplicative Interaction) model depends on that assumption of normally independent distributed error with a constant variance. In the study of genotypes' resistance, disease and pest (insect) incidence on a plant for example, the appropriateness of AMMI model is being doubtful. We can handle it by introducing multiplicative terms for interaction in wider class of modeling, Generalized Linear Models. Its called Generalized AMMI model. An algorithm of iterative alternating generalized regression of row and column estimates its parameters. GAMMI log-link model will be applied to the Poisson data distribution. GAMMI log-link models give us good information of the interaction by its log-odd ratio.

Key words: AMMI, GEI, GAMMI, log-link

INTRODUCTION

The AMMI model represents observations into a systematic component that consists of main effect and interaction effect through multiplication of interactions components, apart from random errors component. Basically, the AMMI analysis combines both additive analysis of variance for the main effect of treatment and analysis of multiple main components uses bilinear modeling for the interaction effect, by using singular value decomposition (SVD) of its interaction matrix (Mattjik & Sumertajaya 2002, Hadi & Sa'diyah 2004, Mattjik 2005). Sometimes, goodness of fit AMMI models which have normally distributed errors with constant variances cannot be satisfied.

Statistical modeling plays the most important role in the providing interpretation of interest phenomenon, and representing it into appropriate language of application field. Transformation can be omitted if homogeneity of variances can be modeled by multiplication of interactions components in the systematic model. However, for non-normally distributed data which is modeled in the observation scale, multiplication of interactions components maybe represent both homogeneity of variances and true multiplication of interactions. It means that there is no warranty that transformation of data in the observation scale be able to separates them.

Transformation, in the regression analysis and analysis of variance cases, has three goals, i.e., to obtain homogeneity of variances, normally of errors, and additional of systemic effects. It is not easy to obtain a satisfaction transformation for all need. So, after transforming, multiplicative component maybe still represents mixture of heterogeneity of variances and multiplicative effects (Hadi et all, 2007).

While, in the additive models, we have widely known generalized linear models (GLM) as a modeling class of non-normally distributed data. In GLM additiveness of systemic effects is given into normally scale. Normally (and homogeneity) of variances is not necessary again. It is because the (quasi) likelihood just need to fix the relationship between mean and variance only.

Multiplicative models (bilinear) bridge the gap between the main effect models (in ANOVA and GLM) and completely interaction models with interaction parameters for each cell in two way table. This models are also give a visually pattern of the main interaction through biplot. Therefore, developing of GLM theory by accommodate the multiplicative component of interaction is very necessary.

The power of multiplicative AMMI model is placed on visualized interactions by using biplot. Van Eeuwijk (1995) proposed the multiplicative model in term of GLM

as an extension of AMMI model called as generalized AMMI (GAMMI) model. In the GAMMI model, it is possibly to visualize interactions. However, as proposed by Van Eeuwijk 1995 that interpretation of the AMMI model must be investigated again, because it depends on link function used in spite of distance its points still represents non-additiveness or dependence. This paper discusses fitting bilinear model in term of GLM, especially for discrete observations, Poisson and Binomial distributions.

GENERALIZED LINEAR MODELS

Classical Linear models have some characteristics, i.e., errors or response variables are normally distributed with constant variances, independent variances of means, and errors or response variables are independent of each other. In the larger model class, it is not depend on these assumptions. Nelder & Wedderburn (1972) discussed GLM which is not depend on these assumptions but it just depends on link function that links between means μ_i and linear predictors η_i of probability distribution model used (Mc Chullagh & Nelder, 1989).

Response variables y_i ($i = 1, 2, \dots, n$) are observation values of random variables Y_i assumed have certain distribution (exponential family) with means $E(y_i) = \mu_i$. In fact, a variance function of mean $V(\mu)$ which accompanies dispersion parameter, follows distribution assumptions $Var(y_i) = \phi V(\mu_i)$ where ϕ is a dispersion parameter (scale factor) and $V(\cdot)$ is variance function. The means μ_i is linked to linear predictors $\eta_i = \sum_{j=1}^n \beta_j x_{ij}$ or $\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta}$ where x_{ij} are known predictor variables, and β_j are unknown parameters to be estimated by using link function $g_i(\mu_i) = \eta_i$. Some distribution variables responses has canonical link function (Table 1). Although, each observation maybe has different link function, it rarely happens. So, the subscript i of function g_i can be omitted or $g_i(\mu_i)$ can be reduced to $g(\mu_i)$. Estimating of parameters β_j in vector $\boldsymbol{\beta}$ can be done by using weighted linear regression iteration procedure of linearized link function which applies to observation (y) of predictor variable (x).

Linearized link functions or working variates (in GLIM), z_i have form $z_i = \eta_i + g(\eta_i)(y_i - \mu_i)$ atau $\mathbf{z} = \boldsymbol{\eta} + (\mathbf{y} - \boldsymbol{\mu}) \delta\boldsymbol{\eta} / \delta\boldsymbol{\mu}$ (Mc Chullagh & Nelder, 1989; Van Eeuwijk, 1995; Falguerolles, 1996). Each observation also has prior weight $w_i = [Var(z_i)]^{-1}$, or $\mathbf{w} = (\delta\boldsymbol{\mu} / \delta\boldsymbol{\eta})^2 / V(\boldsymbol{\mu})$. In every iteration circle, values of x and z are updated. This method is known as iterative reweighted least square (IRLS).

Table 1. Canonical Link functions in generalized linear models

RESPON DISTRIBUTION	NAME	LINK FUNCTION
Normal	Identity	$\eta = g(\mu) = \mu$
Poisson	Log	$\eta = g(\mu) = \log(\mu)$
Binomial	Logit	$\eta = g(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$
Negative Binomial	Log	$\eta = g(\mu) = \log\left(\frac{\mu}{\mu+k}\right)$
Gamma	Invers	$\eta = g(\mu) = \frac{-1}{\mu}$

In general, generalized linear model (GLM) has some characteristics as follows :

- Response variable \mathbf{Y} has distribution in the exponential distribution family;
- Linear or systemic component links linear predictor $\boldsymbol{\eta}$ to multiplication between design \mathbf{X} and parameter $\boldsymbol{\beta}$, $\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta}$;
- Link function $g(\cdot)$ that links linear predictor to fitted values, follows monotonic and differentiable properties. The $g(\cdot)$ describes how the expected response mean is linked to $\boldsymbol{\eta}$, for example, $\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta}$ and $\boldsymbol{\mu} = g^{-1}(\boldsymbol{\eta}) = E(\mathbf{Y})$;
- Response variables may have no constant variances that their values are to be change as changing of their means, $\sigma_i^2 = f(\mu_i)$

GENERALIZED AMMI MODEL

In an experiment, responses observed sometimes is to be categorical data. It affects that AMMI model approach becomes irrelevant. So, it needs to make analysis by using other approach. For this case, AMMI model has also been developed to handle more general cases. This model approach is known as generalized AMMI (GAMMI) model (Van Eeuwijk, 1995) or generalized bilinear models (GBMs) (Falguerles, 1996; Gabriel, 1998). The GAMMI models are given as follows :

$$\eta_{ij} = \mu + \alpha_i + \beta_j + \sum_{k=1}^K \sqrt{\lambda_k} \gamma_{ki} \delta_{kj}$$

An AMMI model is a GAMMI model that its link function is to be identity with constant variance. By letting values of β_j and δ_{kj} , it reduces the model to GLM along the row; while letting values of α_i and γ_{ki} reduces the model to GLM along the column. This GAMMI model's characteristic can be used as base line to

describe parameter estimate procedure. Other parameter estimate procedure usually uses iterative weighted least square.

Fitting GAMMI model is done iteratively through some steps (Van Eeuwijk, 1995; Falguerolles, 1996). The number of multiplicative part in the GAMMI model can be determined by using generalized test of AMMI model as follows :

1. Likelihood ratio test for the first eigen value, and for the second eigen value if the first eigen value has been known, and for the next eigen values. This test compares percentage of those between that can be described by certain part and the total will be described; and it needs no error estimate.
2. F-test needs no special table, and be easy in computation. An independent error estimate (over/under dispersion) is needed and maybe gives problem.
3. Simple test with attribute degree of freedom $(I - 1) + (J - 1) + (2k - 1)$ on the eigen value with respect to pivot k , becomes difference between the number of parameters that will be estimated and the number of applied identification constraints. Appropriate mean square is then tested versus an error estimate (over/under dispersion). This test was proposed by Golob, 1968. When the first eigen value relatively more than the next eigen value, attribute degree of freedom will save to follow Golob, and collect next part to estimate error (over/under dispersion). Sequential application of this procedure is to test eigen value of success versus collected estimate errors.

Adding other multiplicative component for GAMMI model needs recomputing of part that has been inputed. Because of differences cells weight , dimension of success is not nested as usual for the AMMI model with the same cell weight.

Error for diagnostic goal, after convergence, can be obtained from the row regression as well as column regression. Errors of row and column regressions will deviate just a little from the others, because computing of row regression errors assumes column parameters more known than parameters estimate. While for column regression error, estimating of row parameters is also not necessary to be known. Other possibility is to make regressor of parameters result of row and column interactions in the same way as one degree of freedom test for additiveness which gives a regression interpretation, and fits the model with their main effect and regressors. Error of this model is a compromised result between row regression and column regression errors.

Diagnostic error done to assess goodness of fit model is adopted from the GLM. Goodness of fit model can be informally investigated by plotting errors versus fitted values. Generally, for assessment goodness of fit model is suggested to use standardized deviance error and plot it versus linear predictor, or versus fitted

values transformed to information scale constant of error distribution. Transformations of fitted value for some errors distribution are $\hat{\mu}$ for normally distributed error; $2 \sin^{-1} \sqrt{\hat{\mu}}$ for Binomial error; $2 \sqrt{\hat{\mu}}$ for Poisson error; $2 \log(\hat{\mu})$ for Gamma error, and so on.

Goodness of fit model is showed by pattern of errors that spread randomly with constant shift at zero mean. Systemic deviation on this plot is to be in forms: (i) curve form, or (ii) existence of shift change as fitted value change. Curve form can be caused by one of link function use. So, if the plot contain no deviation, we may clime that the link function used is appropriate (model is appropriate). The same thing can also be obtained from errors plot versus linear predictors. Note that this plot is meaningless for binary data. Some other errors plots is used to investigate variance function and link function which are used (Mc Chullagh & Nelder, 1989).

Plot absolute value of error versus fitted value gives informal investigation about properness variance function assumed. This properness is showed by spreading around of points horizontally, and no indication of trends or certain pattern. Improperness of variance function is showed by trend on mean, i.e., positive trend shows that that variance function currently used is gradually increase as increasing of mean and vice versa. Informal investigation about link function used can be done by investigating plot working variate versus linear predictor, but it is not general and meaningless, especially for Binomial distribution.

Visualization Interactions by Using Biplot GAMMI Models

Biplot is very good to show multiplicative interaction in AMMI models. In Biplot, rows and columns are represented by points in two or three dimension spaces. Coordinate of points are obtained from rows and columns scores. Singular values are placed to rows and columns scores in different ways based on what interest, i.e., the relationships among rows, columns, or among rows and columns. By plotting rows scores $v_{ki} = v_{ki} \sqrt{\lambda_k}$, the distances among rows points are proportional to the number of rows interactions. Plotting $\delta_{kj} = \delta_{kj} \sqrt{\lambda_k}$ transforms this relationship to column points. By assuming row and column point as terminal point of vector started at initial point, simple-geometrically can be showed that the number of interactions, or non-additive, between a row and column can be approached by inner-product among vectors in Biplot. This inner-product can be obtained by projecting one of row or column vectors to each other, and then multiplying between length of vector and length of vector where the projection occurs. For the wide class of GAMMI model, it is probable to visualize interaction by using Biplot, but its interpretation depends on certain link function.

Log-Bilinear GAMMI Models

Here we give an other special example of GAMMI model which is to be row x column Goodman’s model (RC Goodman’s model) for two ways frequencies table I x J. The model assumes that each cell I x J are independent and has Poisson distribution. P_{ij} is probability of observations placed in i^{th} row and j^{th} column,

$$P_{ij} = \alpha_i^* \beta_j^* \exp \left(\sum_{k=1}^K \sqrt{\lambda_k} \nu_{ki} \delta_{kj} \right)$$

where α_i^* and β_j^* are positive parameters.

As an identification constrain of multiplicative interaction part, we use same constrain as AMMI model constrain. By taking logarithm, that model is equivalent to log-bilinear model:

$$\eta_{ij} = \log(P_{ij}) = \mu + \alpha_i + \beta_j + \sum_{k=1}^K \sqrt{\lambda_k} \nu_{ki} \delta_{kj}$$

and known as generalized AMMI model with logarithm as link function, where the μ, α_i and β_j are the logarithmic form of α_i^* and β_j^*

Non-independent form is more relevant to association row x column model than non-additive form. Goodman defined two forms of non-independen as follows.

The first is,

$$\omega_{ij} = \log \left(\frac{P_{ij}}{\alpha_i^* \beta_j^*} \right) = \sum_{k=1}^K \sqrt{\lambda_k} \nu_{ki} \delta_{kj}$$

and the second is Log Odds Ratio :

$$\pi_{ij} = \log \left(\frac{P_{ij} P_{st}}{P_{it} P_{sj}} \right) = \sum_{k=1}^K \sqrt{\lambda_k} (\nu_{ki} - \nu_{ks}) (\delta_{kj} - \delta_{kt})$$

It is defined for cells in rows i and s ; and columns j and t . Scaled row parameter $\nu_{ki} = \nu_{ki} \sqrt{\lambda_k}$ can be interpreted as slope of a weighted linear regression of non-independent measure ω_{ij} on column score δ_{kj} :

$$\sum_{j=1}^J \omega_{ij} \delta_{kj} = \nu_{ki}$$

When δ_{kj} are used as coordinate of row points in biplot, squared distance between two row points tends to non-independent between two rows, because

$$\sum_{k=1}^K (v_{ki} - v_{ks})^2 = \sum_{j=1}^J (\omega_{ij} - \omega_{sj})^2$$

The same relationship can be obtain deductively for v_{ki} and δ_{kj} . Therefore, Goodman recommended to visualize from row points and column points by using $v_{ki} = v_{ki}\sqrt{\lambda_k}$ and $\delta_{kj} = \delta_{kj}\sqrt{\lambda_k}$, respectively.

For simultaneously visualize, he recommended to use $v_{ki}^* = v_{ki}\lambda_k^{0.5(1-c)}$ and $\delta_{kj}^* = \delta_{kj}\lambda_k^{0.5c}$, ($0 \leq c \leq 1$), where the choice of c depends on its weighted-point in row or column. The inner-product between row and column points in simultaneous biplot tends to non-independent measure ω_{ij} where γ and δ are scaled to γ^* and δ^* as follows :

$$\omega_{ij} = \log\left(\frac{P_{ij}}{\alpha_i^* \beta_j^*}\right) = \sum_{k=1}^K \sqrt{\lambda_k} v_{ki} \delta_{kj}$$

$$\omega_{ij} = \sum_{k=1}^K v_{ki}^* \delta_{kj}^* = |\gamma_i^*| |\delta_j^*| \cos(\gamma_i^*, \delta_j^*)$$

where γ_i^* and δ_j^* represent as vector of length K . In the same biplot, inner-product between difference of row points and difference of column points tends to log odds ratio :

$$\pi_{ij} = \log\left(\frac{P_{ij} P_{st}}{P_{it} P_{sj}}\right)$$

$$\pi_{ij} = \sum_{k=1}^K \sqrt{\lambda_k} (v_{ki} - v_{ks})(\delta_{kj} - \delta_{kt})$$

$$\pi_{ij} = \sum_{k=1}^K (v_{ki}^* - v_{ks}^*)(\delta_{kj}^* - \delta_{kt}^*)$$

$$\pi_{ij} = |\gamma_i^* - \gamma_s^*| |\delta_j^* - \delta_t^*| \cos(\gamma_i^* - \gamma_s^*, \delta_j^* - \delta_t^*)$$

where γ_i^* , γ_s^* , δ_j^* and δ_t^* are vectors of length K . Simultaneous biplot provides a very good tool to visualize non-independence in two way table of computing analyzed by association row \times column.

For other GAMMI model, interpretation of relationship biplot still must be investigated. In addition, distance between points of one of rows or columns will

indicate some non-additive or non-independent forms. Simultaneously visualize must be interpreted more carefully, but the inner-product of row or column points still tends to non-additive at linear predictor scale.

Especially, for Poisson data case (log-bilinear model), biplot provides two important informations. Firstly, information about dependence between rows or columns which are showed by distance (square) between rows or between columns on biplot. And the second is information about ratio of two probabilities events (odds ratio).

This information is to be geometrical interpretation that uses vector projection properties. Odds is ratio of two probabilities events. From the two ways table genotype x population of pests can be obtained information about ratio of probabilities. We define x_{ij} as i^{th} row and j^{th} column cells values, and $P_{ij} = x_{ij} / \sum_{ij} x_{ij}$ is probability of i^{th} row and j^{th} column events such that we can compute probability ratio of two genotypes, for instance, i^{th} and s^{th} genotypes are attacked by pests, say, j^{th} pest, as P_{ij} / P_{sj} .

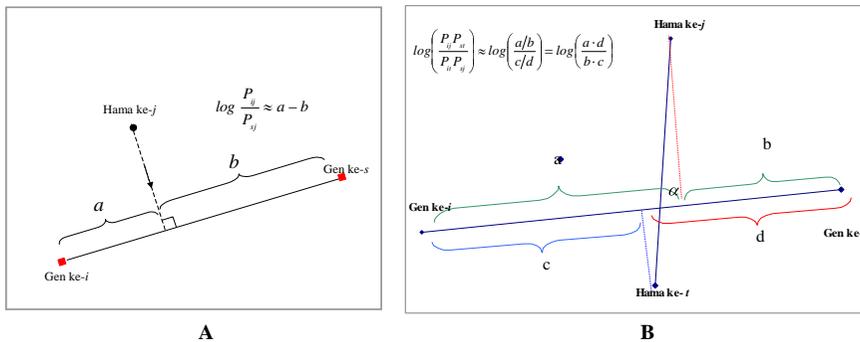


Figure 1. Geometrical visualization of Odds (A) and Odds Ratio (B).

Odds Ratio. Ratio of two odds', for example, ratio between odds of j^{th} pest on i^{th} genotype and s^{th} genotype, and odds of t^{th} pest on same, is given as follows :

$$\frac{P_{ij} / P_{it}}{P_{st} / P_{sj}}$$

This odds ratio can be understood by seeing Figure 1B and considering ratio of difference length between a and b by c and d . Note that if vectors that connect two pests and connect two genotypes are perpendicular each other (ortogonal, $\alpha = 90^\circ$) then this ratio will be equal to 1, and logarithm of it is equal to zero.

According to log-bilinear model (GAMMI log-link) odds ratio can be obtained logarithm scale:

$$\pi_{ij} = \log \left(\frac{P_{ij} P_{st}}{P_{it} P_{sj}} \right) = \sum_{k=1}^K \sqrt{\lambda_k} (v_{ki} - v_{ks})(\delta_{kj} - \delta_{kt})$$

by taking derivation of formula, and gives :

$$\pi_{ij} = \sum_{k=1}^K (v_{ki} - v_{ks})(\delta_{kj} - \delta_{kt})$$

$$\pi_{ij} = |v_i^* - v_s^*| |\delta_j^* - \delta_t^*| \cos(v_i^* - v_s^*, \delta_j^* - \delta_t^*)$$

METODOLOGY

The data is experimental data of leaf pest control on variety of soybean endure that is hybrid result of BALITKABI, Malang, East Java. This experiment involves four varieties of soybean endure resulted by hybrid (Wilis, IAC-100, IAC-80-596-2 and W/80-2-4-20). This research uses leaf pest population data at 14 days old after planting.

Research steps to obtain information about stableness genotype endure are as follows :

1. **Identifying of Distribution and Handling of Experiment data.** Leaf pest population data identified has Poisson distribution. While, percentage of rice-grain and total of grain data identified has binomial distribution. These two sets of data are arranged into two ways table I x J, genotype versus type of leaf pest, and genotype versus location, where cells contain means of replication/block.
2. **Fitting GAMMI Model.** Fitting GAMMI model algorithm is complicated enough, because it is to be a criss-cross regression or alternating regression between row and column regressions, where each regression includes GLM class that is done iteratively by using iterative reweighted least square (IRLS). Therefore, this algorithm involves three types of convergence, i.e., in row regression, in column regression, and in alternating regression. It is complexity of this modeling algorithm. However, basic idea of this algorithm seems easy to be understood.

This algorithm is given in Van Eeuwijk, 1995. Fitting model is done by using software GENSTAT 7th edition or more expand. Poisson distribution data is modeled by using link function logarithm

3. **Analysis of Deviance.** If in AMMI model (ANOVA in general) we test effect of factors by using sum of square then in GAMMI model (GLM in general) we use deviance. We use F-test to determine multiplicative axis/component by

comparing ratio between mean deviance component tested and error mean deviance to F-table value.

4. **Appropriateness Model.** We investigate appropriateness model by using errors diagnostic visually, i.e., errors plot.
5. **Analysis of Stableness Genotype Endure.** Information about stableness genotype endure can be obtained through biplot GAMMI2 configuration. Biplot GAMMI2 gives row and column scores plot (in this case, genotype x pest population or genotype x location) simultaneously. By considering whole biplot, nearness between row and column points shows interaction and dependence (association) of them. Association parameter is given by singular value (generalized). Nearness between certain row point (genotype) and certain column point (pest population or location) indicates association between genotype and certain pest population or location. Small singular value for *i*th GAMMI axis indicates meaningless of the axis.

RESULT AND DISCUSSION

Endurance of Soybean to Leaf Pest. Four genotypes of soybean gave different leaf endurance responses on five types of leaf pest. Table 2 provides population mean of five pests found on four soybean varieties at 14 days old after planting. Based on the algorithm of Eeuwijk, 1995, GAMMI model was using link function natural logarithm and Poisson distribution. Analysis of deviance is given in Table 3.

Table 2. Population means of five types of leaf pest on four genotypes of soybean

Genotype	Types of leaf pest				
	<i>Bemissia</i>	<i>Empoosca</i>	<i>Agromyza</i>	<i>Lamprosema</i>	<i>Longitarsus</i>
IAC-100	0.50	1.75	2.25	0.50	1.75
IAC-80	3.00	2.75	1.00	1.75	3.25
W/80	3.50	4.00	1.25	2.00	2.00
Wilis	4.00	3.00	1.00	1.75	4.00

It shows that the mean of error deviance is 0.0134; and computation of error based on Pearson's Chi-square is 0.0135. Table 3 shows that GAMMI2 model is appropriate because mean deviance ratio of axis 2 is significant at $p\text{-value} < 0.0541 = F_{4,2}$. Singular value of axis 1 and axis 2 are 1.739 and 0.5927, respectively. Plotting of errors deviance versus fitted values and linear predictors shows that there is no anomaly. Plotting of working variate versus linear predictor can indicate improperness link function usage, if the plot is non linear. Also, the plot shows that

there is no anomaly (Figure 3). So that GAMMI2 model which has log-link and has Poisson distribution fits data well.

Biplot GAMMI2 provides information about interaction genotype \times pest (Figure 4). Genotype W/80 shows tendency to be variety candidate that has endurance to all types of leaf pests unless to *Empoasca*, comparing with variety IAC-100 that specifically susceptible to *Agromyza*. Biplot of interaction in log-bilinear model can be used well to find pairs of soybean genotypes and pairs of pest types population which has odds ratio equal to one or log odds ratio equal to zero. On our data, we find that these pairs are genotypes W/80 and IAC-80, and the *Bemisia* and the *Agromyza*. Delivery line of genotypes is “almost” perpendicular to its of pests. It means that log odds ratio “tends” to zero. Table 2 can verify that odds ratio between both of them tends to 1. It means that W/80 and IAC-80 has the same thing, and W/80 and IAC-80 more tend to be attacked by *Bemisia* than the *Agromyza* in the same scale.

Table 3. Analysis of Deviance for leaf pest population data

Source	DF	Deviance	Deviance Mean	Deviance Mean Ratio	p-value
Leaf pest	4	4.1845	1.0461	78.38	0.0126
Genotype	3	2.8359	0.9453	70.83	0.0139
GAMMI 1	6	3.6709	0.6118	45.84	0.0215
GAMMI 2	4	0.9477	0.2369	17.75	0.0541
Error	2	0.0267	0.0133		
Total	19	11.6656	0.6140		

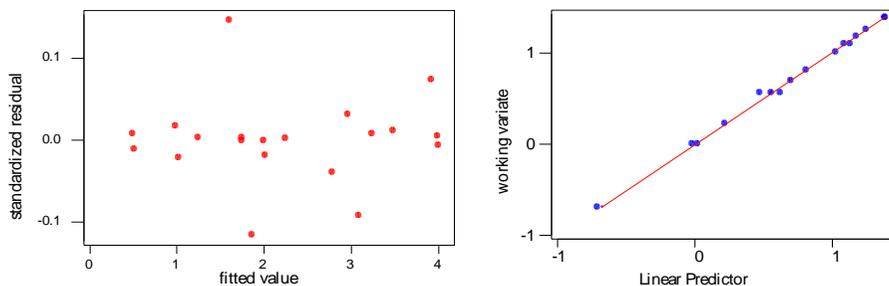


Figure 3. Error plot for soybean pest data: Standardized error versus fitted value GAMMI2 log-link model (left); Working variate versus linear predictor plot (right)

REFERENCES

- Falguerolles, de A, 1996. Generalized Linear-Bilinear Models. An Abstract. Society of Computational Economics. 2nd International Conference on Computing and Finance. Genewa, Switzerland, 26–28 June 1996. <http://www.unige.ch/ce/ce96/defalgue/>
- Gabriel, K. R., 1998, Generalised Bilinear Regression. *Biometrika*. **85** (3):689-700.
- Gollob HF. 1968. A Statistical Model which Combines Features of Factor Analytic and Analysis of Variance Techniques. *Psychometrika* 33:73-116.
- Hadi AF, Sa'diyah H. 2004. AMMI Model untuk Analisis Interaksi Genotipe × Lokasi. *Jurnal Ilmu Dasar* 5(1):33-41
- Hadi AF, Sa'diyah H, Sumertajaya IM. 2007. Penanganan Ketaknormalan Data pada Model AMMI menggunakan Transformasi Box-Cox. *Jurnal Ilmu Dasar* 8(2): 165-174.
- Mattjik A. A. & Sumertajaya I. M. 2002. Perancangan Percobaan dengan Aplikasi SAS dan MINITAB. 2nd Ed. IPB Press. Bogor.
- Mattjik A. A., 2005. Interaksi Genotipe dan Lingkungan dalam Penyediaan Sumberdaya Unggul. Naskah Orasi Ilmiah Guru Besar Biometrika. Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Pertanian Bogor. Bogor.
- McCullagh, P and J. A Nelder 1989. *Generalized Linear Models*. 2nd ed. Chapman and Hall, London.
- Nelder J. A. and R. W. M. Wedderburn. 1972 Generalized Linear Models. *J. Royal Stat. Soc.. Series A (General)*, 135(3): 370-384.
- Van Eeuwijk, F A, 1995. *Multiplicative Interaction in Generalized Linear Models*. *Biometrics*, **51**, 1017–1032

Acknowledgement

1. Prof. Fred van Eeuwijk (University of Wageningen, The Netherland) and Paul Keizer (DLO-Centre for Plant Breeding and Reproduction Research Wageningen, The Netherland) for the discussion on AMMI dan GENSTAT Programming.
2. Dr. Suharsono, reseacher of pathology in soybeen at Indonesian Legumes & Tuber Crop Research Institue, Ministry of Agriculture, The Republik Indonesia.

Halaman ini sengaja dikosongkan

Ketaknormalan Data dan Pengamatan Pencilan pada Model AMMI

Alfian Futuhul Hadi

Mahasiswa Program Doktor Statistika. Sekolah Pascasarjana Institut Pertanian Bogor.
Dosen Jurusan Matematika Universitas Jember
email: afhadi.fmipa@unej.ac.id or afhadi@gmail.com

A. A. Mattjik

Gurubesar Biometrika. Departemen Statistika. Institut Pertanian Bogor.

Abstract. AMMI (Additive Main Effect Multiplicative Interaction) model for interactions in two-way Tabel provide the major mean for studying stability and adaptability through genotype \times environment interaction (GEI), which modeled by full interaction model. Eligibility of AMMI model depends on that assumption of normally independent distributed error with a constant variance. Nowadays, AMMI models have been developed for any condition of MET data which is violence the normality, homogeneity assumption. We can mention in this class of medelling as M-AMMI for mixed AMMI models, G-AMMI for generalized AMMI models. The G-AMMI was handling non-normality i.e categorical response variables using an algorithm of alternating regression. Modeling count data in study of incidence on a plant for example, the appropriateness of AMMI model is being doubtful. GAMMI log-link model will be applied to the Poisson data distribution. GAMMI log-link models give us good information of the interaction by its log-odd ratio. While in handling the non-homogeneity in mix-models sense, one may use a model called factor analytic multiplicative. The development of AMMI models is also to handle any outlier that might be found coincides with non-homogeneity condition of the data. In this paper, we will present both of handling non-normality and outling observation in AMMI model by using an algorithm of alternating regression.

Keywords: AMMI, G-AMMI, M-AMMI, Factor Analytic, Multiplicative Models, Alternating Regression, Robust approach, Biplot.

Pendahuluan

Statistika telah lama berkontribusi terhadap penelitian pemuliaan tanaman, terutama dalam pendekatan biometrika dalam genetika kuantitatif. Lebih jauh, model AMMI telah lama digunakan secara luas untuk menganalisis Interaksi Genotipe \times Lingkungan (IGL) dalam percobaan multi-lingkungan. Hal ini terutama didukung oleh kenyataan bahwa

Model AMMI Kini dan Yang Akan Datang

Pendekatan model AMMI mampu menjelaskan interaksi dengan baik melalui model interaksi lengkap atau dikenal sebagai suku multiplikatif/bilinear (Sumertajaya, 1998). Groenen, P.J.F & A. J. Koning, (2004) menunjukkan penggunaan biplot pada model AMMI sebagai cara baru memvisualisasi interaksi pada model aditif (ANOVA model). Struktur interaksi diuraikan dari matriks sisaan komponen aditif dengan memanfaatkan sifat matematis penguraian nilai singular (singular value decomposition, SVD). SVD merupakan pendekatan kuadrat terkecil dengan reduksi dimensi (pangkat matriks) data yang terbaik dan menyediakan penyajian secara grafis. Model AMMI dibangun dengan landasan teori pemodelan pada distribusi data Normal (gaussian) yang mapan, teknik komputasi yang sederhana, dan telah secara luas digunakan.

Data yang berdistribusi tidak normal cenderung tidak homogen, dan bila dimodelkan dengan AMMI ketakhomogenan ragam dapat berakibat buruk dalam pendugaan. Kadangkala ada alasan kuat untuk tetap memodelkan data pada skala pengamatan, sekalipun skala pengamatan ini cenderung menimbulkan masalah. Kehomogenan ragam dapat diatasi dengan menambahkan satu atau lebih suku multiplikatif interaksi (van Eeuwijk, 1995). Ketika tidak ada alasan untuk memaksa pemodelan tetap pada skala pengamatan, maka transformasi terhadap peubah respon dapat dilakukan untuk mengurangi masalah ini. Model linier atau bilinear kemudian dikenakan pada data tertransformasi, dan sifat sebaran sisaan diasumsikan memenuhi sebaran normal.

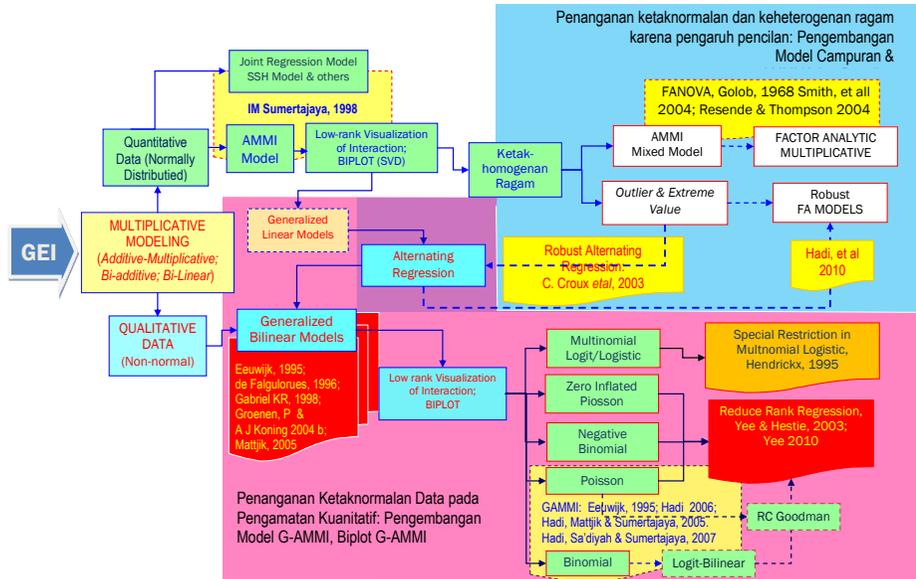
Pada dua dasawarsa terakhir metode seleksi adaptabilitas genotipe melalui percobaan multilokasi telah banyak menggunakan model AMMI. Sampai saat ini model AMMI telah berkembang baik untuk data kuantitatif (sebaran normal), sebut saja model AMMI, maupun untuk data kategorik (sebaran bukan normal) yang disebut model AMMI Terampat atau Generalized AMMI atau disingkat GAMMI.

Namun disisi lain adanya pencilan (pada satu arah tertentu) memungkinkan menjadi pengaruh dalam pengujian ketaknormalan. Dalam hal ini diperlukan kajian sensitifitas pengaruh pencilan terhadap pengujian ketaknormalan. Mendeteksi suatu pengamatan sebagai pencilan dari suatu sebaran normal, atukah pengamatan tersebut menjadi bagian dari suatu sebaran lain yang tidak simetrik (skewed). Hal ini menjadi salah satu kajian yang menarik perhatian penulis. Dalam hal ini, diperlukan metode yang relatif "kekar" terhadap adanya pencilan. Oleh karena itu, pada kesempatan ini penulis memfokuskan pada masalah ketaknormalan data pada pemodelan bilinear dalam dua sebab utama yaitu: (i) karena terdapatnya pengamatan pencilan dan (ii) oleh pencatatan data itu sendiri pada peubah kuantitatif.

Alur (roadmap) Riset

Beberapa konsep pemodelan yang tercakup dalam dua isu tersebut dapat digambarkan secara ringkas melalui gambar 1. Gambar 1 juga menunjukkan fokus utama kajian dalam disertasi ini adalah pada penanganan ketaknormalan oleh (i) adanya pencilan yang disinyalir kerap terjadi bersamaan dengan ketakhomogenan ragam, dan (ii) ketaknormalan oleh sebaran data pengamatn kualitatif. Berdasarkan peta konsep gambar 1, secara teoritik kajian yang dibutuhkan adalah utamanya pada kelas pemodelan multiplikatif, dari

sisi pendugaan parameter, terdapat setidaknya 3 kriteria pendugaan parameter yaitu kuadrat terkecil, simpangan mutlak terkecil dan maksimum likelihood/quasi likelihood. Namun dari sisi komputasi fokusnya adalah pada algoritma regresi bolak-balik (*alternating regression*).



Gambar 1. Peta konsep pemodelan aditif-multiplikatif yang kekar terhadap ketaknormalan dan pengamatan pencilan

Model AMMI dan Asumsi Dasar Kenormalan dan Kehomogenan Ragam

Model AMMI merepresentasikan observasi ke dalam komponen sistematis yang terdiri dari pengaruh utama (main effect) dan pengaruh interaksi melalui suku-suku multiplikatif (multiplicative interactions), disamping komponen acak sisaan atau galat. Komponen acak pada model ini diasumsikan menyebar Normal dengan ragam konstan (Sumertajaya, 1998). Pada dasarnya analisis AMMI menggabungkan analisis ragam aditif bagi pengaruh utama perlakuan dengan analisis komponen utama ganda dengan pemodelan bilinear bagi pengaruh interaksi (Mattjik, 1998, 2005). Struktur interaksi pada model AMMI diuraikan dari matriks sisaan komponen aditif dengan memanfaatkan sifat matematis penguraian nilai singular (singular value decomposition, SVD). SVD merupakan pendekatan kuadrat terkecil dengan reduksi dimensi (pangkat matriks) data yang terbaik dan menyediakan penyajian secara grafis yang dikenal secara luas dengan nama Biplot. Groenen, P.J.F & and Alex J. Koning, (2004b) menunjukkan penggunaan biplot pada model bilinear sebagai cara baru memvisualisasi interaksi. Dalam konteks pemuliaan tanaman --khususnya kajian stabilitas genetik pada komponen daya hasil-- komponen acak pada model ini seringkali diasumsikan berdistribusi normal. Model ini telah berhasil memberikan informasi tentang stabilitas dan adaptasi spesifik suatu genotipe terhadap lingkungan, karena dilengkapi dengan visualisasi matriks IGL melalui Biplot.

Model AMMI Kini dan Yang Akan Datang

Jika matriks data, bersebaran normal dengan ragam konstan, penduga kemungkinan maksimum tereduksi menjadi SVD. Manakala sebarannya bukan normal –Binomial, Poisson, invers Gaussian, misalnya– kesamaan ini tidak lagi berlaku (Falguerolles, 1996). Pada kajian stabilitas ketahanan terhadap penyakit dan kejadian serangan hama pada tanaman misalnya, asumsi model AMMI dengan galat yang normal dan ragam konstan tidak selalu dapat dipenuhi. Pencatatan data populasi hama dalam bentuk cacahan (counting) dan banyaknya buah (polong/gabah) yang terserang penyakit misalnya, merupakan contoh fenomena ini. Lingkungan dalam hal ini adalah lingkungan biotik yang direpresentasi oleh populasi hama/serangan penyakit. Analisis stabilitas dapat digunakan untuk mengidentifikasi ketahanan terhadap hama dan penyakit. Namun kajian IGL menemui kesulitan pada kasus ini (Mattjik, 2005). Data yang berdistribusi selain Normal cenderung tidak homogen, dan bila dimodelkan dengan AMMI ketakhomogenan ragam dapat berakibat buruk.

Model AMMI untuk Kondisi Ketakhomogenan Ragam Akibat Faktor Acak

Model AMMI pada dasarnya adalah model tetap (fixed model) yang mengasumsikan genotipe dan lingkungan ditentukan secara subyektif oleh peneliti dan kesimpulan yang diharapkan hanya terbatas pada genotipe dan lingkungan yang dicobakan saja. Namun telah pula berkembang model campuran, campuran antara faktor tetap dan acak, (Mixed/Random AMMI). Model ini memperluas cakupan kesimpulan, dimana lingkungan bersifat acak dan kesimpulan untuk faktor lingkungan berlaku untuk populasi lingkungan dalam hal ini lokasi budidaya tanaman di seluruh Indonesia.

Mengatasi ketakhomogenan pada AMMI dalam konteks model campuran, dapat digunakan model yang disebut faktor analitik multiplikatif (Smith, *et al*, 2002). Dalam konteks data percobaan multi-lingkungan, pendekatan analisis faktor dapat digunakan untuk menyajikan struktur ragam-peragam dari matriks interaksi dalam sebuah model dengan mempostulat sebuah suku pengaruh genotipe tak-teramati (latent) dalam lingkungan yang berbeda (Smith *et al* 2004). Sedangkan Resende & Thompson, 2004 menyebut postulat itu sebagai ketergatangan pada segugus faktor hipotetik yang bersifat acak.

Dalam sejarah perkembangannya, Gollob 1968 mengenalkan model faktor-analisis ragam (disebut FANOVA) yang menggabungkan aspek analisis ragam dan faktor analisis. Sedangkan Gabriel 1978 mendiskusikan beberapa model yang mirip dengan FANOVA dengan pendugaan parameter menggunakan metode kudrat terkecil. Model AMMI kemudian berkembang dan dikenal sesudah itu. Cornelius, *et al*. 1996 menuliskan bahwa Gaugh & Zobel memberi nama lain pada model FANOVA yang diusulkan Gollob sebagai model AMMI. Smith, 2002 menyimpulkan bahwa model faktor analitik ekuivalen dengan AMMI model campuran. Pengembangan model AMMI juga dilakukan untuk menangani pencilan yang seringkali muncul secara bersamaan dengan kondisi ketakhomogenan ragam dalam data.

Penggunaan Transformasi Kenormalan Data pada Model AMMI

Suatu percobaan multilokasi membutuhkan biaya yang tidak sedikit, disamping melibatkan begitu banyak genotipe, ia juga dilakukan pada berbagai lingkungan budidaya di Indonesia. Desain percobaan di tiap-tiap lokasi haruslah sama satu sama lain. Beberapa asumsi dan kaidah perancangan percobaan adakalanya tidak dapat dipenuhi, seperti kebebasan antar pengamatan, termasuk didalamnya asumsi sebaran data. Beberapa masalah akan muncul dalam analisis seleksi adaptabilitas melalui model AMMI, terutama pada masalah ketaknormalan. Kadangkala ada alasan kuat untuk tetap memodelkan data pada skala pengamatan, sekalipun skala pengamatan ini cenderung menimbulkan masalah. Keheterogenan ragam dapat diatasi dengan menambahkan satu atau lebih suku multiplikatif interaksi (van Eeuwijk, 1995). Ketika tidak ada alasan untuk memaksa pemodelan tetap pada skala pengamatan, maka transformasi terhadap peubah respon dapat dilakukan untuk mengurangi masalah ini.

Transformasi, dalam kasus analisis regresi ataupun analisis ragam, bertujuan untuk memperoleh kehomogenan ragam, mendekati kenormalan galat, dan keaditifan pengaruh sistematis. Tidaklah mudah memperoleh sebuah transformasi yang memenuhi semua kebutuhan. Jadi, setelah transformasi pun, suku multiplikatif kemungkinan masih mencerminkan campuran keheterogenan ragam dan pengaruh multiplikatif. Beberapa penerapan transformasi kenormalan pada model AMMI menggunakan transformasi pangkat Box-Cox dapat dijumpai pada Hadi, *et al* 2007.

Seperti dilaporkan Hadi *et al* 2007, transformasi Box-Cox mampu mengatasi ketaknormalan data. Dengan transformasi Box-Cox dapat dilakukan pemodelan interaksi menggunakan model AMMI secara sah pada data ternormalkan. Namun transformasi kenormalan dilakukan semata-mata untuk tetap bertahan pada model dengan metode pendugaannya yang telah mapan secara teori sehingga pengujian hipotesis dan interpretasinya pun tidak banyak perdebatan. Model AMMI dibangun dengan landasan teori pemodelan yang mapan, teknik komputasi yang sederhana, dan telah secara luas digunakan. Transformasi data dilakukan untuk semata-mata memperoleh asumsi kenormalan. Analisis AMMI kemudian dilakukan pada data hasil transformasi ini. Kita seolah menutup mata terhadap makna apa yang diberikan oleh transformasi pada interpretasi model AMMI.

Perluasan Model AMMI untuk Data Berdistribusi Taknormal

Pada pemodelan aditif telah dikenal luas apa yang disebut dengan Generalized Linear Models (GLM) sebuah kelas pemodelan yang menangani data-data non-normal. Model ini tidak bergantung pada karakteristik atau asumsi distribusi data (kenormalan dan kehomogenan ragam), tetapi bergantung hanya pada sifat fungsi hubung (link function) yang menghubungkan parameter rata-rata dan prediktor linear (linear predictor) dari model sebaran peluang yang digunakan. Namun model ini merupakan model aditif,

Model AMMI Kini dan Yang Akan Datang

sampai dengan tahun 1995, belum mengadopsi model interaksi lengkap seperti pada AMMI.

Eeuwijk, 1995 dan Falguolues, 1996 mengusulkan algoritma regresi bolak-balik untuk masing-masing faktor interaksi. Namun seperti dilaporkan Hadi *et al* 2008 algoritma ini diperkirakan akan menemui kompleksitas karena didalamnya akan terlibat sedikitnya 3 kekonvergenan. Di samping itu, algoritma ini menemui masalah utama ketakkonstanan ragam.

Model multiplikatif (bilinear) menjembatani kesenjangan antara model pengaruh utama (pada ANOVA atau GLM) dan model interaksi lengkap dengan satu parameter interaksi untuk tiap-tiap sel dalam tabel dua arah. Model ini pun memberikan visualisasi pola utama interaksi melalui Biplot. Karenanya pengembangan teori GLM dengan mengakomodasi komponen multiplikatif untuk interaksi sangat diperlukan. Introduksi model multiplikatif untuk kelas pemodelan GLM, sebagai generalisasi dari AMMI, disebut sebagai Generalized Linear Additive Main-Effect & Multiplicative Interactions Model (Model GAMMI).

Kekuatan model multiplikatif AMMI terletak pada visualisasi interaksi melalui Biplot. Pada model GAMMI, visualisasi interaksi ini masih dimungkinkan. Namun masih harus diinvestigasi karena sangat tergantung pada distribusi data dan link-function yang digunakan. Hadi, AF 2008 melaporkan bahwa link-function ini mempengaruhi interpretasi pada perbandingan peluang (odds) dan rasio odds yang dibutuhkan pada interpretasi.

Karenanya pengembangan teori MLT dengan mengakomodasi komponen multiplikatif untuk interaksi sangat diperlukan. Sejauh ini beberapa penggunaan model AMMI terampat pada distribusi poisson dan binomial telah dilakukan sebagaimana oleh Hadi *et al* 2007, Hadi *et al* 2008. Tingkat serangan penyakit yang dicatat dalam rendah, sedang, dan tinggi misalnya, atau pencatatan kualitatif pada pengamatan intensitas warna dan tekstur buah merupakan contoh pengamatan multinomial. Perluasan dari G-AMMI diperlukan untuk menangani sebaran lain seperti multinomial dan ordinal.

Antisipasi Pengaruh Pencilan pada Model AMMI

Pengamatan pencilan sering mendapat perhatian pada analisis statistika. Ukuran pemusatan dan penyebaran data yang sering digunakan seperti rata-rata mempunyai sifat rentan terhadap adanya pencilan. Akibat satu pengamatan yang sangat jauh berbeda nilainya dari sesamanya akan cenderung mempengaruhi nilai rata-rata. Demikian juga dengan ragam, karena ragam diukur menggunakan konsep rata-rata jarak kuadrat pengamatan terhadap rata-ratanya maka ia juga rentan terhadap pencilan. Hampir semua prosedur statistik yang mendasarkan pada statistik rata-rata dan ragam dan sebaran normal pada umumnya, secara teori akan menghadapi hal yang sama. Termasuk di dalamnya pendekatan kuadrat (jarak) terkecil atau Least Square.

Pencilan (pada satu arah tertentu) memungkinkan menjadi pengaruh dalam pengujian ketaknormalan. Dalam hal ini diperlukan kajian sensitifitas pengaruh pencilan terhadap pengujian ketaknormalan. Mendeteksi suatu pengamatan sebagai pencilan dari suatu sebaran normal atau sebaran simetrik, ataukah pengamatan tersebut menjadi bagian dari suatu sebaran tertentu yang tidak simetrik.

Pada data numerik yang mengikuti sebaran normal, keberadaan pencilan pada satu sisi tertentu akan menimbulkan kemenjuluran (skewness) yang cenderung terdeteksi sebagai ketaknormalan. Model AMMI sebagaimana model-model lain yang menggunakan Singular Value Dekomposition (SVD) seperti Analisis Komponen Utama dan Analisis Faktor, rentan terhadap adanya pencilan, karena SVD berbasis Least Square. Sedangkan dalam upaya merakit sifat-sifat unggul pada suatu genotipe, pencilan justru menjadi sesuatu yang berharga, karenanya mengabaikannya tidaklah bijaksana. Untuk itu diperlukan metode yang relatif "kekar" terhadap adanya pencilan.

Selain rata-rata dan ragam, dikenal pula ukuran pemusatan dan penyebaran lain yaitu median dan simpangan mutlak median atau Median Absolute Deviation (MAD). Dua statistik ini berbasis median yang relatif kekar terhadap pencilan. Beberapa prosedur statistika juga dikembangkan menurut keduanya, antara lain regresi kekar dengan konsep simpangan median mutlak terkecil atau least absolute menggantikan least square. Prosedur-prosedur ini kemudian disebut dalam kelas pemodelan atau pendekatan "robust".

Dalam kasus penerapan model AMMI untuk analisis IGL dan kestabilan, meski belum pernah dilaporkan sebelumnya, namun karena secara teoritik SVD berbasis least square, maka tetap berpotensi menghadapi masalah akibat pengamatan pencilan. Untuk membangun kekekaran pada model AMMI maka dasarnya adalah mengubah kriteria pemodelan dari least square ke least absolute, baik itu pada pendugaan pengaruh utama dan pada pengaruh interaksi. Cukup sulit untuk membayangkan bagaimana analisis ragam yang sudah begitu melekat dalam kehidupan ilmiah harus kita ubah menjadi analisis simpangan mutlak atau istilah lain bagi pendekatan robust. Betapapun sulit dan tidak populer, hal tersebut tetap harus dilakukan sebagai suatu langkah ilmiah yang didasari pemikiran visioner. Mungkin jalan ke arah itu masih jauh. Sementara itu, pemecahan jangka pendek adalah memanfaatkan algoritma pada model AMMI terampat yaitu regresi bolak-balik terampat. Sebagaimana disebutkan sebelumnya bahwa model AMMI Terampat dengan sebaran data normal dan fungsi hubungan identitas adalah ekuivalen dengan model AMMI klasik. Merakit sifat kekekaran terhadap pencilan pada model AMMI dapat dilakukan dengan strategi migrasi dari pendekatan regresi bolak-balik dengan basis least square ke regresi bolak-balik dengan regresi least absolute. Hal inilah yang dilakukan oleh Hadi, dan Mattjik 2009b.

Sifat Kekar Pencilan

Sebagaimana lazimnya, matriks data Y berukuran $n \times p$ digunakan untuk melambangkan catatan yang berisi peubah pengamatan/karakteristik (pada kolom) atas beberapa individu (objek/case) pada barisnya. Dengan peubah-peubah tersebut yang sudah

Model AMMI Kini dan Yang Akan Datang

distandardisasi sehingga mempunyai nilai tengah nol dan ragam 1. Suatu skor faktor dicatat sebagai f_i . Dengan vektor skor ke- i diberikan sebagai $f_i = (f_{i1}, \dots, f_{ik})^T$, sedangkan loading vektor ke- j adalah $\lambda_j = (\lambda_{j1}, \dots, \lambda_{jk})^T$ keduanya tidak diketahui. Vektor dari seluruh loading dan skor faktor dilambangkan dengan $\theta = (f_1^T, \dots, f_n^T, \lambda_1^T, \dots, \lambda_p^T)$, dan

$$\hat{y}_{ij}(\theta) = \sum_{l=1}^k f_{il} \lambda_{lj} = f_i^T \lambda_j = \lambda_j^T f_i$$

adalah nilai dugaan (fitted value) bagi y_{ij} . Dengan memilih θ sedemikian sehingga nilai dugaan dan nilai aktual dari matriks sangat dekat, kita mendefinisikan nilai dugaan vektor skor \hat{f}_i dan $\hat{\lambda}_j$ sebagai nilai dugaan vektor loading. Nilai dugaan matriks data \hat{Y} dapat didekomposisi menjadi $\hat{Y} = \hat{F} \hat{\Lambda}^T$ dengan baris-baris dari \hat{F} adalah nilai dugaan skor dan baris-baris-baris dari $\hat{\Lambda}^T$ adalah nilai dugaan loadings.

Perhatikan bahwa *rank* dari \hat{Y} sebesar-besarnya adalah $k < p$, sedangkan *rank* dari Y adalah sebesar p . Pendekatan metode kuadrat terkecil (least squares, selanjutnya disebut LS) adalah meminimumkan jumlah kuadrat sisaan:

$$\hat{\theta}_{LS} = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^n \sum_{j=1}^p (y_{ij} - \hat{y}_{ij}(\theta))^2$$

dan menghasilkan nilai dugaan \hat{Y} yang dapat dilihat sebagai pendekatan "terbaik" (dalam kacamata kuadrat terkecil) terhadap matriks data Y oleh matriks ber-*rank* k . Teorema Eckart-Young (Gower and Hand 1996, p. 241) menyebutkan bahwa aproksimasi terbaik ini dapat diperoleh melalui dekomposisi nilai singular (singular value decomposition/SVD) terhadap matriks data $Y = UDV^T$. Dengan mengganti seluruh nilai singular dalam matriks D dengan nol, kecuali untuk k buah nilai singular terbesar pertama, dapat diperoleh D_k dan akhirnya $\hat{Y} = UD_kV^T$. Dan dengan menyatakan $\hat{F} = \sqrt{n}U$ dan $\hat{\Lambda} = VD_k/\sqrt{n}$ kita memperoleh apa yang disebut solusi komponen utama untuk masalah faktor analisis (cfr. Johnson and Wichern 1998, p. 524). Selanjutnya, matriks peragam contoh dari vektor skor dugaan adalah $\hat{F}^T \hat{F} = I_k$ dimana hal ini konsisten dengan asumsi awal $\operatorname{Cov}(F) = I_k$.

Penting untuk kita catat bahwa nilai dugaan \hat{F} dan $\hat{\Lambda}^T$ hanya dinyatakan secara khusus untuk suatu transformasi linier. Dan karena $\hat{Y} = (\hat{F}^T)^T (\hat{\Lambda}^T)^T$ untuk sembarang matriks T non-singular berukuran $k \times k$ matrix, maka $\hat{F}^T T$ and $\hat{\Lambda}^T T^{-1}$ mendapat nilai yang sama bagi fungsi tujuan $\hat{\theta}_{LS}$. Meski demikian nilai dugaan \hat{Y} yang diperoleh bersifat unik. Selanjutnya, jika kita menambahkan kendala bahwa dugaan matriks peragam bagi vektor skor adalah matriks satuan, maka dugaan \hat{F} dan $\hat{\Lambda}^T$ menjadi suatu transformasi orthogonal, sehingga matrix $\hat{\Lambda} \hat{\Lambda}^T$ juga unik.

Karena kriteria kuadrat terkecil memberikan pembonot yang berlebih kepada siaan yang bernilai besar, gagasan awla mengatasinya adalah menggunakan kriteriasimpangan multak terkecil (Least Absolute Deviations) atau disebut kriteria L1, yang mana diketahui memberikan penduga aditif bagi tabel dua arah yang robust (Terbeck and Davies 1998).

Kriteria L1 ini menghasilkan penduga yang meminimumkan simpangan mutlak sebagai berikut $\hat{\theta}_{L1} = \operatorname{argmin}_{\theta} \sum_{i=1}^n \cdot \sum_{j=1}^p |y_{ij} - \hat{y}_{ij}(\theta)|$

Untuk memperoleh \hat{F} dan $\hat{\Lambda}^T$ yang optimal, haruslah dipenuhi bahwa nilai dugaan \hat{f}_i meminimumkan $\sum_{j=1}^p |y_{ij} - f_i^T \hat{\lambda}_j|$ dan $\hat{\lambda}_j$ meminimumkan $\sum_{i=1}^n |y_{ij} - \hat{f}_i^T \lambda_j|$

Oleh karena itu, pendekatan ini tidaklah meminimumkan kedua jumlah mutlak pada kriteria L1 secara bersamaan, melainkan dengan menetapkan indeks j dan skor f_i kemudian mencari λ_j yang meminimumkan $\sum_{i=1}^n |y_{ij} - f_i^T \lambda_j|$. Masalah ini sekarang bukan lagi masalah bilinear tetapi regresi linier dengan algoritma simpangan mutlak terkecil.

Dan segera dapat kita lihat bahwa meminimumkan $\sum_{i=1}^n |y_{ij} - f_i^T \lambda_j|$ dilakukan satu persatu untuk setiap $j = 1, \dots, p$ berkenaan dengan kriteria L1 untuk skor yang ditetapkan. Analog dengan itu, untuk loading λ_j yang ditetapkan, mencari f_i yang meminimumkan $\sum_{j=1}^p |y_{ij} - f_i^T \lambda_j|$ (untuk setiap $i = 1, \dots, n$ bergantian) berkenaan dengan meminimumkan kriteria L1 pada nilai loadings yang diberikan. Bergantian bolak-balik pada meminimuman $\sum_{i=1}^n |y_{ij} - \hat{f}_i^T \lambda_j|$ dan $\sum_{j=1}^p |y_{ij} - \hat{f}_i^T \lambda_j|$ membawa kita pada skema pendugaan iteratif yang disebut alternating regression (regresi bolak-balik). Catatlah bahwa nilai kriteria L1 menurun pada setiap langkah iteratif itu.

Algoritma yang serupa, tetapi berbasis pada regresi kuadrat terkecil klasik yang bolak-balik, dilakukan oleh de Falguerolles and Francis 1992, sedangkan Gabriel 1998 menggunakan regresi bolak-balik dengan basa linier terampat (generalized linear models) untuk model bilinear terampat (generalized bilinear models).

Sayangnya, regresi L1 sensitif terhadap pengamatan pengungkit (leverage points). Jika terdapat skor atau loading yang memencil maka regresi L1 dapat sangat terpengaruh olehnya. Dengan memberikan pembobot yang bernilai lebih rendah (downweighting) terhadap pengamatan pengungkit ini kita mendapatkan regresi L1 terboboti (weighted L1 regression), dan menghasilkan penduga sebagai berikut :

$$\hat{\theta}_{RAR} = \operatorname{argmin}_{\theta} \sum_{i=1}^n \cdot \sum_{j=1}^p w_i(\theta) v_j(\theta) |y_{ij} - \hat{y}_{ij}(\theta)|$$

Satu fungsi objektif tunggal menduga \hat{F} dan $\hat{\Lambda}^T$ secara simultan dari baris-baris dan kolom-kolom matriks Y . Untuk mendapatkan nilai dugaan $\hat{\theta}_{RAR}$ digunakan regresi bolak-balik robust (Robust Alternating Regressions, disingkat RAR). Penduga $\hat{\theta}_{RAR}$ tidak akan salah arah (mislead) oleh adanya pengamatan pencilan.

Pembobot bagi baris, $w_i(\theta)$ didefinisikan oleh $w_i(\theta) = \min(1, \frac{\chi_{k,0.95}^2}{RD_i^2})$ untuk $i = 1, \dots, n$ dimana $\chi_{k,0.95}^2$ adalah batas atas bagi nilai kritis 5% sebaran khi-kuadrat dengan derajat bebas k , dan $RD_i = \sqrt{(f_i - T(F))^T C(F)^{-1} (f_i - T(F))}$ untuk $i = 1, \dots, n$ adalah jarak kekar (robust distances) yang dihitung dari koleksi vector skor $F = \{f_i | 1 \leq i \leq n\}$ pada ruang berdimensi k (Rousseeuw and van Zomeren 1990). Dan analog dengan itu, segugus

pembobot bagi kolom v_j didefinisikan dengan cara yang sama menggunakan vektor loading. Catatlah bahwa karena loading dan skor yang sebenarnya tidak teramati, w_i dan v_j bergantung pada vektor parameter θ yang tidak diketahui.

Ilustrasi 1: Memodelkan Data Cacahan pada Percobaan Ketahanan Hama Daun Kedelai

Dari percobaan pengendalian terhadap hama daun pada galur kedelai tahan hasil persilangan oleh Balitkabi di Malang, Jawa Timur. Percobaan ini melibatkan empat galur/varietas kedelai tahan hasil persilangan (Wilis, IAC-100, IAC-80-596-2 dan W/80-2-4-20). Penelitian ini memanfaatkan data populasi hama daun pada umur 14 hari setelah tanam. Keempat genotipe kedelai memberikan respon ketahanan daun yang berbeda terhadap lima jenis hama daun. Tabel 1 menyajikan rata-rata populasi kelima hama yang ditemui pada keempat varietas kedelai pada usia 14 hari setelah tanam. Dengan algoritma bolak-balik dalam Hadi *et al* 2007, model GAMMI menggunakan fungsi hubungan logaritma natural dan sebaran Poisson. Analisis devians disajikan pada Tabel 2 menunjukkan bahwa rata-rata residual devians adalah 0.0134; pada perhitungan sisaan berbasis Khi-kuadrat Pearson sebesar 0.0135. Tabel 2 menunjukkan bahwa model GAMMI-2 memenuhi kelayakan, karena rasio rata-rata devians sumbu 2 signifikan pada nilai- $p < 0.0541$ F-tabel [4,2]. Nilai singular sumbu 1 dan 2 berturut-turut adalah 1.739, 0.5927.

Tabel 1. Rataan populasi lima jenis hama daun pada empat genotipe kedelai

Genotipe	Jenis Hama Daun				
	<i>Bemissia</i>	<i>Empoasca</i>	<i>Agromyza</i>	<i>Lamprosema</i>	<i>Longitarsus</i>
IAC-100	0.50	1.75	2.25	0.50	1.75
IAC-80	3.00	2.75	1.00	1.75	3.25
W/80	3.50	4.00	1.25	2.00	2.00
Wilis	4.00	3.00	1.00	1.75	4.00

Tabel 2. Analisis devians untuk data populasi hama daun

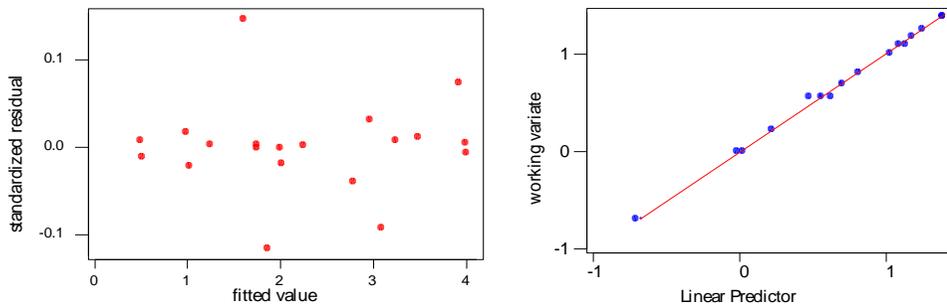
Sumber	Derajat Bebas	Devians	Rataan Devians	Rasio Rataan Devians	Nilai-p
Hama Daun	4	4.1845	1.0461	78.38	0.0126
Genotipe	3	2.8359	0.9453	70.83	0.0139
GAMMI 1	6	3.6709	0.6118	45.84	0.0215
GAMMI 2	4	0.9477	0.2369	17.75	0.0541
Residual	2	0.0267	0.0133		
Total	19	11.6656	0.6140		

Plot residual devians terhadap nilai dugaan model dan linear prediktor, menunjukkan tidak adanya kelainan yang berarti. Plot antara working variate terhadap prediktor linier

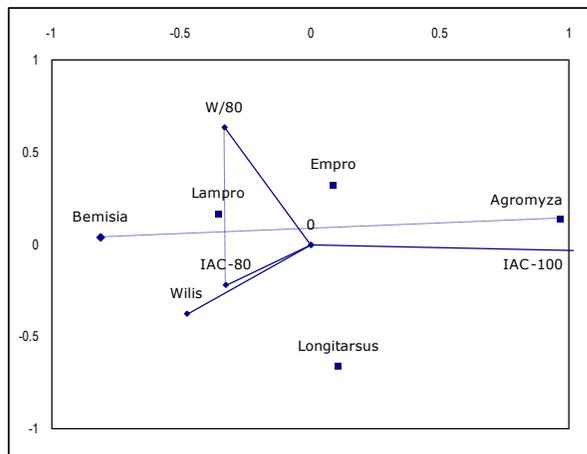
Model AMMI Kini dan Yang Akan Datang

dapat mengindikasikan ketidaktepatan penggunaan fungsi hubung, jika plot ini tidak linier. Tidak ada penyimpangan pada plot ini (Gambar 3). Sehingga model GAMMI-2 dengan log-link dan distribusi Poisson tampak mengepas data dengan baik.

Biplot GAMMI-2 menyajikan informasi interaksi genotipe \times hama. Genotipe W/80 tampak berpeluang untuk menjadi kandidat varietas yang relatif tahan terhadap semua jenis hama daun kecuali pada *Empoasca*, itupun hanya jika dibandingkan dengan varietas IAC-100 yang secara spesifik rentan terhadap *Agromyza* (Gambar 3). Biplot interaksi model log-bilinear dapat digunakan secara baik untuk menemukan pasangan genotipe kedelai dan pasangan populasi jenis hama yang mempunyai rasio odds satu atau log-rasio odds nol. Pada data kita, ditemui bahwa pasangan itu adalah genotipe W/80 dan IAC-80 terhadap hama Bemisia dan Lalat. Garis antar genotipe "hampir" tegak lurus dengan garis antar jenis hama menunjukkan log-rasio odds "mendekati" nol.



Gambar 2. Plot residual untuk data hama kedelai: Plot residual terstandarisasi terhadap nilai dugaan model GAMMI-2 *log-link* (kiri); Plot *working variate* terhadap prediktor linier (kanan).



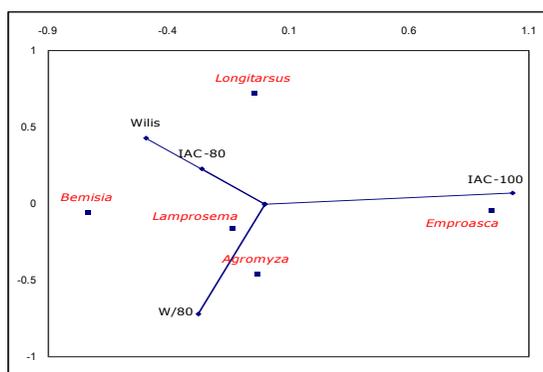
Gambar 3. Biplot GAMMI-2 untuk interaksi hama daun dengan fungsi hubung logaritma.

Tabel 1 dapat memverifikasi bahwa rasio odds antara keduanya mendekati 1. Artinya W/80 dan IAC-80 mempunyai kesamaan, W/80 cenderung terserang Bemisia daripada Lalat, demikian pula dengan IAC-80 dalam skala (odd rasio) yang sama.

Model AMMI Kini dan Yang Akan Datang

Tabel 3. Analisis ragam untuk populasi hama daun yang ditransformasi

Sumber	Derjat Bebas	Jumlah Kuadrat	Kuadrat Tengah	Pengujian 1 Suku Multiplikatif		Pengujian 2 Suku Multiplikatif	
				Nilai F	Nilai-p	Nilai F	Nilai-p
Hama Daun	4	2.2659	0.5665	5.64	0.032	50.02	0.0197
Genotipe	3	1.6252	0.5417	5.40	0.039	47.83	0.0205
AMMI 1	6	2.1065	0.3511	3.50	0.077	31.00	0.0316
AMMI 2	4	0.5797	0.1449			12.80	0.0738
Residual	2	0.0227	0.0113				
Total	19	6.5999	0.3474				



Gambar 4. Biplot AMMI 2 data populasi hama daun yang ditransformasi dengan Box-Cox

Model AMMI pada data hama daun yang ditransformasi dengan pangkat 0.66 melalui metode transformasi box-cox menghasilkan matriks interaksi sebagai berikut:

$$\begin{pmatrix} -0.527851 & -0.062761 & 0.665758 & -0.109114 & 0.033889 \\ 0.116444 & -0.130709 & -0.184108 & -0.023206 & 0.221626 \\ 0.192549 & 0.441590 & -0.134987 & 0.176737 & -0.675950 \\ 0.218851 & -0.248173 & -0.346669 & -0.044436 & 0.420517 \end{pmatrix}$$

Sedangkan model GAMMI Log-link data hama daun menghasilkan matriks interaksi sebagai berikut:

$$\begin{pmatrix} 0.532780 & 0.037319 & 0.624311 & -0.241579 & 0.112729 \\ 0.139683 & -0.136263 & -0.234720 & 0.004502 & 0.226799 \\ 0.194269 & 0.326580 & -0.036568 & 0.246332 & -0.730612 \\ 0.198829 & -0.227636 & -0.353022 & -0.009255 & 0.391085 \end{pmatrix}$$

Perbandingan kedua matriks interaksi yang dihasilkan kedua metode ini menggunakan metode procrustes diperoleh nilai R kuadrat sebesar 98.73%. Angka ini menunjukkan

bahwa pada pendugaan matriks interaksi kedua metode ini sangat dekat, tidak banyak berbeda. Apakah ini berasal dari peran penggunaan transformasi Box-Cox, tidaklah sertamerta kita dapat katakan demikian. Sebab bila kita menggunakan AMMI secara langsung pada data asal tanpa transformasi diperoleh R-kuadrat procrustes sebesar 98.26%. Hal yang paling mungkin berperan dalam hal ini adalah karakter distribusi data rata-rata populasi hama mirip dengan sebaran Normal.

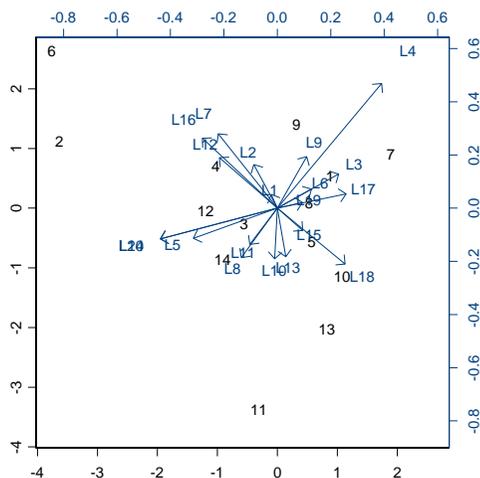
Bila kita perhatikan sel baris pertama kolom pertama kedua pada kedua matriks tersebut di atas, terlihat angka yang sama cukup besar (dibandingkan angka pada sel-sel lain) namun berbeda tanda. Secara geometris hal ini berarti pada dimensi tersebut titik ini berada pada posisi yang berlawanan sehingga menyebabkan perbedaan pada konfigurasi kedua matriks ini. Namun bila titik-titik lain relatif sama maka perbedaan ini menjadi tidak tampak atau tidak terdeteksi oleh metode procrustes, karena secara matematis metode procrustes tidak memperhatikan "tanda". Karena itulah meskipun kedua matriks interaksi di atas sangat mirip, namun interpretasi kestabilan/ketahanan terhadap hama penyakit dapat saja berbeda.

Pada pemodelan terampat, GAMMI atau GLM secara umum, transformasi bukanlah tidak dilakukan. Justru transformasi menjadi hal terpenting dalam model linier terampat. Transformasi dalam model linier terampat dipilih berdasarkan distribusi data, karenanya identifikasi distribusi menjadi sangat penting. Transformasi inilah yang memodelkan data berdasarkan fungsi peluangnya melalui fungsi hubung, dan secara teori interpretasi pemodelan inipun diturunkan berdasarkan fungsi peluang ini. Dengan bilinear terampat (GAMMI), kita sedang memodelkan peluang, sehingga interpretasinya berkenaan dengan perbandingan peluang, odds dan rasio odds. Namun di sisi lain, transformasi ini juga yang menyebabkan dibutuhkan metode pendugaan parameter yang lain (*likelihood*) bukan pendugaan dengan Metode Kuadrat Terkecil (MKT) biasa. Konsekuensi lain, model bilinear terampat khususnya, membutuhkan komputasi yang tidak sesederhana biasanya.

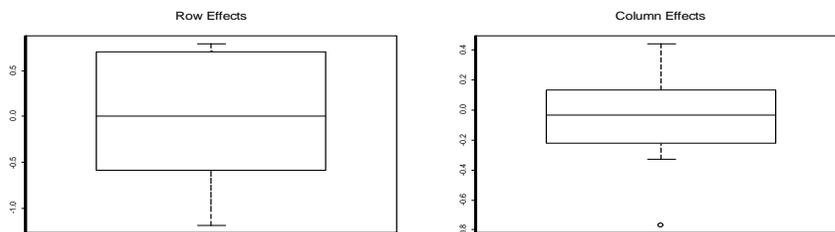
Bila kita memilih transformasi kenormalan sebagai upaya penanganan data-data bukan Normal pada pemodelan bilinear, kita dapat menggunakan transformasi Box-Cox dan kemudian memodelkannya dengan AMMI, bila distribusi data yang kita miliki mendekati simetrik, sangat mirip dengan sebaran Normal. Hal ini disebabkan karena hasil AMMI dengan pendekatan transformasi Box-Cox tidak jauh berbeda dengan penggunaan GAMMI. Namun bila distribusi data bukan Normal kedua pendekatan ini mungkin memberikan hasil yang sangat berbeda.

Yang paling penting adalah, informasi tentang rasio odds pada model log-bilinear tidak dapat diperoleh pada model AMMI dari data ternormalkan. Ini menjadi kelebihan model GAMMI *log-link* dibanding model AMMI dengan transformasi kenormalan pada data berdistribusi Poisson. Demikianlah model AMMI terampat (GAMMI) mampu memberikan informasi sebagaimana model AMMI, dengan keluasan distribusi peubah respon.

Ilustrasi 2: Mengepas Model AMMI Kekar Pencilan



Gambar 5. Biplot IGL Model FANOVA Kekar Pencilan (factor 1 49.43%, cumulative 2 factor 53.76%)



Gambar 6. Boxplot pengaruh baris (Genotipe) dan kolom (Lingkungan) data konsorsium padi

Data dari Konsorsium Padi Nasional yang dihimpun oleh Balai Besar Padi Sukamandi. Percobaan yang dilakukan pada tahun 2008 ini bertujuan untuk mengevaluasi performa galur padi pada 20 lingkungan budidaya yang berbeda. Melibatkan 11 genotipe yang terdiri dari 3 genotipe dari BATAN, 4 dari BB Padi, 2 dari Biogen, dan 2 dari IPB, dengan 3 varietas pembanding yaitu Ciherang, Inpari1, dan Cimelati. Kita akan menggunakan matriks IGL yang berupa tabel dua arah yang selnya berisi median dari 3 ulangan percobaan. Dengan algoritma sebagaimana Hadi AF, & AA Matjik 2009b diperoleh hasil yang menunjukkan bahwa tidak terdapat pencilan yang “membahayakan” dalam data. Hanya satu pencilan terdeteksi pada pengaruh kolom (lingkungan) sebagaimana pada Gambar 6. Sehingga dapat dipastikan hasilnya akan sangat mirip dengan pendekatan least-square. Dari gambar 6 dapat dikatakan bahwa lingkungan 4 (L4) memiliki “keragaman” yang relatif besar. Genotipe-genotipe yang dekat dengan titik asal adalah genotipe yang relatif stabil yaitu G3, G8, G4, G12, and G14. Ketika tidak ada pencilan sama sekali, maka hasil dari pendekatan robust ini akan sama dengan yang diperoleh dari pendekatan least square. Gambar 6 menunjukkan bahwa hanya satu pengamatan yang

ditengarai sebagai pencilan pada pengaruh kolom yaitu pada L4 dan padanya telah diberikan bobot yang rendah dalam perhitungan.

Model FANOVA Gollob 1968 menggabungkan aspek analisis ragam dan analisis faktor, pendugaan parameternya menggunakan kriteria least square. Model AMMI kemudian dikembangkan sesudah itu. bahwa Gauch & Zobel memberi nama lain pada model FANOVA yang diusulkan Gollob sebagai model AMMI. Ketika kita menangani ketakhomogenan ragam dengan model campuran, kita dapat memodelkan data menggunakan model faktor analitik. Model faktor analitik ekuivalen dengan model AMMI. Keduanya menggunakan prinsip dekomposisi metriks interaksi yang berbasis kriteria least square. Pada dasarnya ide merakit sifat kekekaran terhadap pencilan ini adalah memberikan bobot yang lebih rendah terhadap pengamatan pencilan dan kemudian menduga parameternya dengan meminimumkan kesalahan dengan kendala atau kriteria yang konsisten dengan pendekatan robust yaitu *least absolute*.

Pustaka

- Cornelius, P.L., J. Crossa, M.S, Seyedsadr. 1996, Statistikal Test and Estimators of Multiplicative Models for GEI. In Genotype by Environment Interaction, Manjit S. Kang, & HG. Gauch (eds), CRC Press. (pp. 200) <http://books.google.co.id/books?id=NPqjC-mvRTUC&lpg=PP1&pg=PA199> Visit at 10 July 2009.
- Croux C. and Haesbroeck G. 1999. Influence function and efficiency of the minimum covariance determinant scatter matrix estimator, *Journal of Multivariate Analysis*, 71:161-190.
- Croux C. and Haesbroeck G. 2000. Principal component analysis based on robust estimators of the covariance or correlation matrix: influence functions and efficiencies, *Biometrika*, 87:603-618.
- de Falguerolles, A, 1996. Generalized Linear-Bilinear Models. An Abstract. Society of Computational Economics. 2nd International Conference on Computing and Finance. Genewa, Switzerland, 26–28 June 1996. <http://www.unige.ch/ce/ce96/defalgue/>
- Gauch H.G. 1988. Model selection and validation for yield trial with interaction, *Biometrics*, 44:705-716.
- Gauch, H.G., Jr (1992) *Statistikal Analysis of Reginal Yield Traials: AMMI Analysis of Factorial Designs*. Elsevier, Amsterdam, The Netherlands.
- Gollob H.F. 1968. A statistikal model which combines features of factor analytic and analysis of variance techniques, *Psychometrika*, 33:73-116.
- Gabriel, K. R., 1998, *Generalised Bilinear Regression*. *Biometrika*. 85 (3):689-700.
- Greenacre, M. J. 1984. *Theory and Aplications of Correspondence Analysis*. Academic Press. London.
- Jolliffe, I T. 1986. *Principal Component Analysis*. Springer-Verlag. New York.
- Hadi, A. F. & H. Sa'diyah, 2004. AMMI Model untuk Analisis Interaksi Genotipe × Lokasi. *Jurnal Ilmu Dasar*1:33-41
- Hadi A. F., A. A. Mattjik, I. M. Sumertajaya, & H. Sa'diyah. 2008. AMMI Model on Count Data: Model Log-Bilinier. The 3rd International Conference on Mathematics and Statistiks (ICoMS-3). Moslem Statistikian and Mathematician in South East Asia (MSMSEA). Bogor, Indonesia, 5-6 August 2008

Model AMMI Kini dan Yang Akan Datang

- Hadi, A. F. & A. A. Mattjik, 2009a. Generalized AMMI Models For Assessing The Endurance of Soybean To Leaf Pest. The First International Seminar on Science and Technology (ISSTEC 2009). Universitas Islam Indonesia, Yogyakarta, 24 – 25 January 2009.
- Hadi, A.F, & A.A. Mattjik 2009b. Developing Robustness Of The AMMI Models By Robust Alternating Regression. Proceeding at The 4rd International Conference on Mathematics and Statistiks. Bandar Lampung, August 2009.
- Hadi, A. F., H. Sa'diyah, & I. M. Sumertajaya. 2007. Handling Non-normality Data on AMMI Models using BoxCox Transformation. Jurnal Ilmu Dasar, FMIPA Universitas Jember. Vol 8. No 2. pp. 165-174.
- Hadi A. F, I. M. Sumertajaya, & I.M. Tirta, 2008. AMMI Model on Binomial Data: Model Logit-Bilinear. National Seminar Nasional on Science & Tech. University of Lampung, Bandar Lampung 17-18 November 2008.
- Mattjik A. A. & Sumertajaya I. M. 2002. Perancangan Percobaan dengan Aplikasi SAS dan MINITAB. 2nd Ed. IPB Press. Bogor.
- Mattjik A. A., 2005. Interaksi Genotipe dan Lingkungan dalam Penyediaan Sumberdaya Unggul. Naskah Orasi Ilmiah Guru Besar Biometrika. Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Pertanian Bogor. Bogor.
- McCullagh, P. and Nelder, J. A. 1989. Generalized Linear Models. 2nd ed. Chapman and Hall, London.
- Sumertajaya, I M. 1998. Perbandingan Model AMMI dan Regresi Linier untuk Menerangkan Pengaruh Interaksi Percobaan Lokasi Ganda. Tesis. Program Studi Statistika Sekolah Pascasarjana IPB, Bogor
- Resende, M. D. P & R. Thompson, 2004. Factor Analytic Multiplicative Mixed Model in The Analysis of Multiple Experiments. Rev. Mat. Estat., vol. 22. No. 2. Sao Paulo.
- Rousseeuw P.J. and Van Driessen K. 1999. A fast algorithm for the minimum covariance determinant estimator, Technometrics, 41:212-223.
- Smith. A., B. Cullis, D. Lockett, G. Hollamby and R. Thompson. 2002. Exploring genotype-environment data using random effects AMMI models with adjustments for spatial field trend: Part 1-2: Theory-Applications. In. Kang, M.S. (Ed). Quantitative Genetics, Genomics and Plant Breeding. CAB International, UK
- Terbeck W. and Davies P. 1998. Interactions and outliers in the two-way analysis of variance, The Annals of Statistiks, 26:1279-1305.
- Ukkelberg A. and Borgen O. 1993. Outlier detection by robust alternating regressions, Analytica Chimica Acta, 277:489-494.
- Van Eeuwijk, F A, 1995. *Multiplicative Interaction in Generalized Linear Models. Biometrics*, **51**, 1017–1032

Ucapan Terima Kasih

1. Tulisan ini bagian dari penelitian HIBAH PENELITIAN PASCASARJANA, No. 41/13.24.4/SPK/BG-PD/2009, 30 Maret 2009.
2. Dr. Ir. Aan Darajat, Peneliti Senior Balai Besar Padi dan Dr Suharsono, Balitabi, Malang atas ijin penggunaan data penelitiannya.

Halaman ini sengaja dikosongkan

10

KOMPUTASI BAYESSIAN PADA MODEL AMMI

10.1 Bayessian

Terdapat dua paradigma utama yang digunakan dalam melakukan pendugaan parameter, yaitu *Frequentist* dan *Bayesian*. Perbedaan utama dari dua paradigma tersebut terletak pada informasi yang digunakan dalam pendugaan parameter. Pada paradigma *Frequentist*, parameter diasumsikan bernilai tetap dan pendugaan parameternya hanya didasarkan pada informasi yang dibawa oleh contoh, sedangkan pada paradigma *Bayesian* parameter dari model yang akan diduga diasumsikan memiliki

sebaran yang bersifat acak. Selain itu, informasi yang digunakan dalam proses pendugaan berasal dari informasi yang diperoleh dari contoh dan informasi awal mengenai parameter yang akan diduga.

Fungsi kepekatan dari informasi awal disebut sebagai fungsi kepekatan awal (sebaran prior) dari θ ($\pi(\theta)$). Sedangkan fungsi kepekatan peubah acak X dipandang sebagai fungsi kepekatan bersyarat $X|\theta$ yang ditulis sebagai $f(x|\theta)$. Sementara $f(x,\theta)$ merupakan notasi dari fungsi kepekatan bersama X dan θ yang didefinisikan sebagai $f(x,\theta) = f(x|\theta) \pi(\theta)$. Sementara X memiliki fungsi kepekatan marginal:

$$m(x) = \int_{\theta} f(x|\theta) dF\pi(\theta), \text{ untuk } \theta \text{ peubah acak kontinu}$$

Untuk $m(x) > 0$ dapat diperoleh sebaran posterior sebagai berikut:

$$\pi(\theta|x) = \frac{f(x,\theta)}{m(x)}$$

Fungsi $\pi(\theta|x)$ dinamakan sebagai sebaran posterior yang didefinisikan sebagai sebaran bersyarat θ jika data contoh x diketahui. Sebaran posterior merupakan suatu sebaran yang dibentuk dari hasil kombinasi antara informasi awal dan informasi dari contoh. Sebaran posterior merupakan sebaran dasar pengujian dalam metode Bayesian (Berger, 1985).

Salah satu kendala penerapan metode Bayesian untuk pendugaan parameter adalah penentuan sebaran prior. Sebaran prior merefleksikan pengetahuan atau keyakinan peneliti tentang parameter. Dalam menentukan sebaran prior seringkali mempertimbangkan kemudahan dalam membuat sebaran posterior, karena secara umum tidak mudah menghitung $m(x)$ dan $\pi(\theta|x)$ (Berger, 1985). Kelas sebaran prior yang membuat sebaran posterior dapat ditentukan dengan mudah karena posterior memiliki keluarga sebaran yang sama dengan keluarga sebaran prior disebut sebagai *conjugate prior*.

Masalah utama dalam penerapan pendekatan Bayesian terletak pada cara mendapatkan sebaran posterior yang sering memerlukan proses pengintegralan fungsi yang berdimensi tinggi. Hal ini dapat menyebabkan perhitungan menjadi sulit. *Markov Chain Monte Carlo* (MCMC) adalah suatu metode yang dapat digunakan untuk tujuan tersebut.

Ada dua bagian pengertian dari MCMC yaitu “*Monte Carlo*” yang berhubungan dengan proses simulasi secara acak dan “*Markov Chain*” yang berhubungan dengan proses sampling suatu nilai baru dengan syarat nilai sebelumnya dari sebaran posterior (Lynch, 2007). Algoritma MCMC diimplementasikan dengan cara mengambil contoh berulang-ulang dari p sebaran posterior bersyarat $[\vartheta_1/\vartheta_2, \dots, \vartheta_p], \dots, [\vartheta_p/\vartheta_1, \dots, \vartheta_{p-1}]$ (Albert, 2009).

Gibbs sampling merupakan salah satu metode untuk membangun algoritma *Markov Chain Monte Carlo* (MCMC). *Gibbs sampling* adalah suatu teknik untuk membangkitkan peubah acak dari sebaran (marjinal) secara tidak langsung, tanpa perlu menghitung fungsi kepekatannya (Casella & George, 1992).

Gibbs Sampling dapat diterapkan apabila sebaran peluang bersama (*joint probability distribution*) tidak diketahui secara eksplisit, tetapi sebaran bersyarat (*conditional distribution*) dari tiap-tiap variabel diketahui (Hoff, 2009). Misalkan diketahui suatu vector dari parameter $\theta = \{\theta_1, \theta_2, \dots, \theta_p\}$, dan informasi mengenai ukuran peluang θ adalah $p(\theta) = p(\theta_1, \theta_2, \dots, \theta_p)$. Dengan memberi nilai awal $\theta^{(0)} = \{\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_p^{(0)}\}$, Gibbs sampling akan membangkitkan $\theta^{(l)}$ dari $\theta^{(l-1)}$ seperti berikut.

a. Untuk $l=1, 2, \dots, m$, dibangkitkan:

1. contoh $\theta_1^{(l)} \sim p(\theta_1 | \theta_2^{(l-1)}, \dots, \theta_p^{(l-1)})$
2. contoh $\theta_2^{(l)} \sim p(\theta_2 | \theta_1^{(l)}, \theta_3^{(l-1)}, \dots, \theta_p^{(l-1)})$
- :
- :
3. contoh $\theta_p^{(l)} \sim p(\theta_p | \theta_1^{(l)}, \theta_2^{(l)}, \dots, \theta_{p-1}^{(l)})$

b. dilakukan proses yang sama sampai $l = m$ yang menunjukkan proses sudah konvergen. Dua dari beberapa literatur yang membahas mengenai pengujian kekonvergenan antara lain Ntzoufras (2009) dan Megersen at all. (1998).

10.2 Pendekatan Bayesian pada Pendugaan Matriks Interaksi

Percobaan yang melibatkan dua faktor telah umum digunakan pada uji daya hasil tanaman. Uji daya hasil atau disebut juga uji multilokasi bertujuan untuk meningkatkan keakuratan pendugaan daya hasil, melihat kestabilan hasil dan pola respon genotif antar lingkungan, serta membantu pemulia menentukan genotif-genotif terbaik untuk direkomendasikan sebagai varietas baru. Dua faktor yang biasanya dilibatkan dalam uji daya hasil yakni faktor lingkungan dan genotif.

Model linier untuk uji multilokasi dengan genotif dan lingkungan sebagai perlakuan adalah sebagai berikut:

$$y_{ijk} = \mu + \tau_i + \rho_{k(j)} + \gamma_j + \delta_{ij} + \varepsilon_{ijk}$$

dengan:

y_{ijk} = respon dari genotif ke- i pada lingkungan ke- j dalam kelompok ke- k

μ = nilai rata-rata umum

τ_i = pengaruh genotif ke- i , $i=1,2,\dots,a$

$\rho_{k(j)}$ = pengaruh kelompok ke- k tersarang pada lingkungan ke- j , $k=1,2,\dots,r$

γ_j = pengaruh lingkungan ke- j , $j=1,2,\dots,b$

δ_{ij} = pengaruh interaksi genotif ke- i dengan lingkungan ke- j

ε_{ijk} = pengaruh sisaan dari genotif ke- i dalam kelompok ke- k yang dilakukan di lingkungan ke- j

Jika parameter dari model linear di atas akan diduga melalui pendekatan komputasi Bayesian, maka diperlukan sebaran prior dari masing-masing parameter. Sebaran prior pada parameter di semua lingkungan didefinisikan sebagai sebaran normal dengan nilai tengah nol dan ragam sesuai dengan kondisi yang diinginkan (Edwards and Jannink, 2006). Sebaran prior berikut yang digunakan untuk komputasi dengan pendekatan Bayesian pada model dengan interaksi:

$$\mu \sim N(\mu_\mu, \sigma_\mu^2);$$

$$\tau \sim N(\mu_\tau, \sigma_\tau^2);$$

$$\gamma \sim N(\mu_\gamma, \sigma_\gamma^2);$$

$$\delta_{ij} \sim N(\mu_\delta, \sigma_{\delta_{ij}}^2);$$

$$\sigma^2 \sim IG(\alpha_\sigma, \beta_\sigma)$$

Dengan mengkombinasikan antara informasi awal (sebaran prior) dengan informasi tentang parameter tersebut yang dibawa oleh data observasi akan diperoleh sebaran posterior dari masing-masing parameter.

Secara analitik, fungsi sebaran posterior diperoleh dari perkalian antara prior dengan likelihood.

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

Sebaran untuk $(Y_{ijk} | \theta)$ adalah:

$$Y_{ijk} | \theta \sim N(\eta_{ij}, \sigma^2)$$

dengan

$$\eta_{ij} = \mu + \tau_i + \gamma_j + \delta_{ij}$$

dan θ didefinisikan sebagai

$$\mu, \tau_i, \gamma_j, \delta_{ij}, \sigma^2$$

Sehingga didapat Likelihoodnya sebagai berikut:

$$\begin{aligned} L(\theta) &= \prod_{ijk} (\pi \sigma^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2} (Y_{ijk} - \eta_{ij})^2\right\} \\ &= (\pi \sigma^2)^{-abr/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{ijk} (Y_{ijk} - \bar{y}_{ij})^2 - \frac{r}{2\sigma^2} \sum_{ij} (Y_{ij} - \eta_{ij})^2\right\} \end{aligned}$$

Sebaran posterior bersamanya adalah:

$$\begin{aligned} \pi(\theta | y^n) &\propto L(\theta) \times \pi_\mu(\mu) \times \pi_\tau(\tau) \times \pi_\gamma(\gamma) \times \pi_\delta(\delta) \times \pi_{\sigma^2}(\sigma^2) \\ \pi(\theta | y^n) &\propto L(\theta) \times \pi_\mu(\mu) \times \pi_\tau(\tau) \times \pi_\gamma(\gamma) \times \pi_\delta(\delta) \times \pi_{\sigma^2}(\sigma^2) \\ &= (\pi \sigma^2)^{-abr/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{ijk} (Y_{ijk} - \bar{y}_{ij})^2 - \frac{r}{2\sigma^2} \sum_{ij} (Y_{ij} - \eta_{ij})^2\right\} \\ &\quad \times (\pi \sigma_\mu^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_\mu^2} (\mu - \mu_\mu)^2\right\} \times (\pi \sigma_{\tau_i}^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_{\tau_i}^2} (\tau_i - \mu_{\tau_i})^2\right\} \\ &\quad \times (\pi \sigma_{\gamma_j}^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_{\gamma_j}^2} (\gamma_j - \mu_{\gamma_j})^2\right\} \times (\pi \sigma_{\delta_{ij}}^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_{\delta_{ij}}^2} (\delta_{ij} - \mu_{\delta_{ij}})^2\right\} \\ &\quad \times \frac{\beta^\alpha}{\Gamma(\alpha)} \sigma^{-\alpha-1} \exp\left(-\beta/\sigma\right) \end{aligned}$$

Sebaran posterior dari masing-masing parameter diperoleh dari perkalian antara prior dari parameter dengan likelihood.

- Sebaran posterior untuk μ (Liu, 2001)

$$\begin{aligned} \pi(\mu | \theta \text{ lainnya}) &\propto \exp\left\{-\frac{r}{2\sigma^2} \sum_{ij} (\hat{y}_{ij} - \eta_{ij})^2\right\} \times \exp\left\{-\frac{1}{2\sigma_\mu^2} (\mu - \mu_\mu)^2\right\} \\ &\propto \exp\left\{-\frac{r}{2\sigma^2} \sum_{ij} (\hat{y}_{ij} - \mu)^2\right\} \times \exp\left\{-\frac{1}{2\sigma_\mu^2} (\mu - \mu_\mu)^2\right\} \\ &\propto \exp\left\{-\frac{rab\sigma_\mu^2 + \sigma^2}{2\sigma^2\sigma_\mu^2} \left(\mu - \frac{rab\sigma_\mu^2\bar{y}_{\dots} + \sigma^2\mu_\mu}{rab\sigma_\mu^2 + \sigma^2}\right)^2\right\} \end{aligned}$$

$$\mu | \theta \text{ lainnya} \sim N\left(\frac{rab\sigma_\mu^2\bar{y}_{\dots} + \sigma^2\mu_\mu}{rab\sigma_\mu^2 + \sigma^2}, \frac{\sigma^2\sigma_\mu^2}{rab\sigma_\mu^2 + \sigma^2}\right)$$

- Sebaran posterior untuk τ_i

$$\begin{aligned} \pi(\tau_i | \theta \text{ lainnya}) &\propto \exp\left\{-\frac{r}{2\sigma^2} \sum_{ij} (\hat{y}_{ij} - \eta_{ij})^2\right\} \times \exp\left\{-\frac{1}{2\sigma_{\tau_i}^2} (\tau_i - \mu_{\tau_i})^2\right\} \\ &\propto \exp\left\{-\frac{rb\sigma_{\tau_i}^2 + \sigma^2}{2\sigma^2\sigma_{\tau_i}^2} \left(\tau_i - \frac{rb\sigma_{\tau_i}^2\hat{\tau}_i + \sigma^2\mu_{\tau_i}}{rb\sigma_{\tau_i}^2 + \sigma^2}\right)^2\right\} \end{aligned}$$

$$\tau_i | \theta \text{ lainnya} \sim N\left(\frac{rb\sigma_{\tau_i}^2\hat{\tau}_i + \sigma^2\mu_{\tau_i}}{rb\sigma_{\tau_i}^2 + \sigma^2}, \frac{\sigma^2\sigma_{\tau_i}^2}{rb\sigma_{\tau_i}^2 + \sigma^2}\right)$$

- Sebaran posterior untuk γ_j

$$\begin{aligned} \pi(\gamma_j | \theta \text{ lainnya}) &\propto \exp\left\{-\frac{r}{2\sigma^2} \sum_{ij} (\hat{y}_{ij} - \eta_{ij})^2\right\} \times \exp\left\{-\frac{1}{2\sigma_{\gamma_j}^2} (\gamma_j - \mu_{\gamma_j})^2\right\} \\ &\propto \exp\left\{-\frac{ra\sigma_{\gamma_j}^2 + \sigma^2}{2\sigma^2\sigma_{\gamma_j}^2} \left(\gamma_j - \frac{ra\sigma_{\gamma_j}^2\hat{\gamma}_j + \sigma^2\mu_{\gamma_j}}{ra\sigma_{\gamma_j}^2 + \sigma^2}\right)^2\right\} \end{aligned}$$

$$\gamma_j | \theta \text{ lainnya} \sim N\left(\frac{ra\sigma_{\gamma_j}^2\hat{\gamma}_j + \sigma^2\mu_{\gamma_j}}{ra\sigma_{\gamma_j}^2 + \sigma^2}, \frac{\sigma^2\sigma_{\gamma_j}^2}{ra\sigma_{\gamma_j}^2 + \sigma^2}\right)$$

- Sebaran posterior untuk δ_{ij}

$$\pi(\delta_{ij} | \mu, \tau_i, \gamma_j, \sigma^2) \propto \exp\left\{-\frac{r}{2\sigma^2} \sum_{ij} (\delta_{ij} - \eta_{ij})^2\right\} \times \exp\left\{-\frac{1}{2\sigma_{\delta_{ij}}^2} (\delta_{ij} - \mu_{\delta_{ij}})^2\right\}$$

$$\propto \exp\left\{-\frac{r\sigma_{\delta_{ij}}^2 + \sigma^2}{2\sigma^2\sigma_{\delta_{ij}}^2} \left(\delta_{ij} - \frac{r\sigma_{\delta_{ij}}^2 \hat{\delta}_{ij} + \sigma^2 \mu_{\delta_{ij}}}{r\sigma_{\delta_{ij}}^2 + \sigma^2}\right)^2\right\}$$

$$\delta_{ij} | \theta \text{ lainnya} \sim N\left(\frac{r\sigma_{\delta_{ij}}^2 \hat{\delta}_{ij} + \sigma^2 \mu_{\delta_{ij}}}{r\sigma_{\delta_{ij}}^2 + \sigma^2}, \frac{\sigma^2 \sigma_{\delta_{ij}}^2}{r\sigma_{\delta_{ij}}^2 + \sigma^2}\right)$$

- Sebaran posterior untuk σ^2

$$\pi(\sigma^2 | \theta \text{ lainnya}) \propto L(\theta) \times \pi_{\sigma^2}(\sigma^2 | \alpha_{\sigma}, \beta_{\sigma})$$

$$\propto (\pi_{\sigma^2})^{abr/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{ijk} (\delta_{ijk} - \eta_{ij})^2\right\} \times (\sigma^2)^{-(\alpha_{\sigma}+1)} \exp\left\{-\frac{\beta_{\sigma}}{\sigma^2}\right\}$$

$$\propto (\sigma^2)^{-(abr/2+\alpha_{\sigma}+1)} \exp\left\{-\frac{1}{\sigma^2} \left(\beta_{\sigma} + \frac{1}{2} \sum_{ijk} (\delta_{ijk} - \eta_{ij})^2\right)\right\}$$

$$\sigma^2 | \theta \text{ lainnya} \sim IG\left[\frac{abr}{2} + \alpha_{\sigma}, \beta_{\sigma} + \frac{1}{2} \sum_{ijk} (\delta_{ijk} - \eta_{ij})^2\right]$$

Untuk menduga parameter model melalui pendekatan Bayesian selanjutnya dihitung dengan *Gibbs sampling*. Proses yang dilakukan pada tahap ini adalah (Silvianti, 2009):

1. Dihitung nilai dugaan awal dari parameter model menggunakan metode kuadrat terkecil (MKT):

- a. $\hat{\mu} = \frac{1}{abr} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r y_{ijk}$

- b. $\hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..}$

- c. $\hat{\gamma}_j = \bar{y}_{.j} - \bar{y}_{..}$

- d. $\hat{\delta}_{ij} = \bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$

- e. Penguraian nilai singular dari $\hat{\delta}_{ij} = VLS^T$, $\hat{\lambda}_k$ merupakan unsur diagonal dari L, \hat{v}_k merupakan vektor dari kolom matriks V, dan \hat{s}_k merupakan vektor dari kolom matriks S.

- f. $\hat{\mu} = \frac{1}{abr} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r y_{ijk}$

- g. $\hat{\tau}_i = \mu_i - \mu$

$$h. \hat{\gamma}_j = \mu_j - \mu$$

$$i. \hat{\delta}_{ij} = \mu_{ij} - \mu - \tau_i - \gamma_j$$

2. Dibangkitkan sebaran posterior untuk $\theta=(\hat{\mu}, \hat{\tau}_i, \hat{\gamma}_j, \hat{\delta}_{ij})$ menggunakan *Gibbs sampling*. Nilai awal dari setiap parameter menggunakan nilai dugaan awal yang diperoleh seperti pada tahap 1.
3. Nilai rata-rata dari sebaran posterior yang dibangkitkan pada tahap 2 digunakan untuk menduga μ, τ, γ , dan δ , dimana:

$$a. \tilde{\mu} = \frac{1}{m} \sum_{l=1}^m \mu^{(l)}$$

$$b. \tilde{\tau}_i = \frac{1}{m} \sum_{l=1}^m \tau_i^{(l)}$$

$$c. \tilde{\gamma}_j = \frac{1}{m} \sum_{l=1}^m \gamma_j^{(l)}$$

$$d. \tilde{\delta}_{ij} = \frac{1}{m} \sum_{l=1}^m \delta_{ij}^{(l)}$$

Berdasarkan tahapan di atas, nilai dugaan dari pengaruh interaksi δ_{ij} menggunakan pendekatan komputasi Bayesian adalah

$$\tilde{\delta}_{ij} = \frac{1}{m} \sum_{l=1}^m \delta_{ij}^{(l)}$$

10.3 Model AMMI dengan Komputasi Bayessian (G N A Wibawa, Aunuddin, A A Mattjik)

Pada model AMMI, pengaruh interaksi dari genotif dan lingkungan (δ_{ij}) diuraikan menjadi komponen utama interaksi (KUI). dengan demikian, model AMMI dapat dituliskan dalam bentuk persamaan sebagai berikut:

$$y_{ijk} = \mu + \tau_i + \rho_{k(j)} + \gamma_j + \sum_{n=1}^m \sqrt{\lambda_n} v_{in} s_{jn} + \phi_{ij} + \varepsilon_{ijk}$$

dengan:

- y_{ijk} = respon dari genotif ke-i pada lingkungan ke-j dalam kelompok ke-k
- μ = nilai rata-rata umum
- τ_i = pengaruh genotif ke-i, $i=1,2,\dots,a$
- $\rho_{k(j)}$ = pengaruh kelompok ke-k tersarang pada lingkungan ke-j, $k=1,2,\dots,r$
- γ_j = pengaruh lingkungan ke-j, $j=1,2,\dots,b$
- $\sqrt{\lambda_n}$ = nilai singular untuk komponen bilinear ke-n, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$
- v_{in} = pengaruh ganda genotif ke-i melalui komponen bilinear ke-n
- s_{jn} = pengaruh ganda lingkungan ke-j melalui komponen bilinear ke-n
- ϕ_{ij} = sisaan dari komponen bilinear
- ε_{ijk} = pengaruh sisaan dari genotif ke-i dalam kelompok ke-k yang dilakukan di lingkungan ke-j
- m = banyaknya KUI yang dipertahankan dalam model

Dengan pendekatan klasik, KUI diduga menggunakan penguraian nilai singular (*singular value decomposition*). Sedangkan jika KUI diduga menggunakan pendekatan Bayesian, maka diperlukan sebaran prior dari setiap parameter dari KUI (λ , v , dan s). Penggunaan komputasi Bayesian telah digunakan oleh Viele & Srinivasan (1999) untuk menduga parameter model AMMI pada data dengan ukuran contoh tidak sama dan ragam heterogen. Liu (2001) mengembangkan pendekatan ini untuk menduga semua parameter model AMMI dan mengecek kekonvergenan sebaran serta kekonsistenan pemilihan model.

Seperti pada pendugaan matriks interaksi dengan pendekatan Bayesian, sebaran prior yang digunakan dari setiap parameter merupakan *conjugate prior*. Untuk memperoleh *conjugate prior* diperlukan sebaran dari setiap parameter. Sebaran dari akar ciri adalah normal positif, sementara vektor ciri menyebar menurut sebaran von-Mises Fisher (Viele & Srinivasan,1999; Liu,2001).

Misalkan $X \sim N(\mu, \sigma^2)$, suatu peubah acak Y dikatakan menyebar normal positif (N^+) jika sebaran dari Y proporsional terhadap sebaran X untuk $y \geq 0$, dan 0 untuk y lainnya.

Suatu vektor satuan acak \mathbf{x} ($\|\mathbf{x}\|=1$) berdimensi p dikatakan menyebar menurut sebaran von-Mises Fisher, $M_p(\mu, k)$, jika memiliki fungsi kepekatan peluang (Mardia dan Jupp, 2000; Dillon dan Sra, 2003):

$$M_p(\mu, k) = c_p(k)e^{k\mu^T \mathbf{x}}, \mathbf{x} \in S^{p-1}$$

Dengan $\|\mu\|=1$, $k \geq 0$, S^{p-1} adalah unit hypersphere berdimensi p, dan $c_p(k)$ adalah

$$c_p(k) = \frac{k^{\frac{p}{2}-1}}{(2\pi)^{\frac{p}{2}} I_{\frac{p}{2}-1}(k)}$$

$I_p(k)$ merupakan fungsi Bessel yang dimodifikasi pada ordo ke-p

$$I_p(k) = \sum_{k \geq 0} \frac{1}{\Gamma(p+k+1)k!} \left(\frac{k}{2}\right)^{2k+p}$$

dengan $\Gamma(\cdot)$ merupakan fungsi Gamma.

Parameter μ menunjukkan rata-rata arah dan k menunjukkan *concentration parameter*. Jika $k=0$, maka \mathbf{x} menyebar menurut sebaran seragam sperikal (Mardia dan Jupp, 2000).

Sebaran von-Mises Fisher merupakan sebaran keluarga eksponensial (Mardia dan El-Atoum, 1976; Nuñez-antonio dan Gutiérrez-peña, 2005). *Conjugate prior* dari sebaran ini juga merupakan sebaran von-Mises Fisher.

Prior yang digunakan untuk menduga parameter model AMMI dengan komputasi Bayesian adalah *conjugate prior* yaitu (Viele & Srinivasan, 1999; Liu, 2001):

$$\mu \sim N(\mu_\mu, \sigma_\mu^2); \quad \tau \sim N(\mu_\tau, \sigma_\tau^2); \quad \gamma \sim N(\mu_\gamma, \sigma_\gamma^2); \quad \sigma^2 \sim IG(\alpha_\sigma, \beta_\sigma); \\ \lambda_n \sim N^+(\mu_\lambda, \sigma_\lambda^2); \quad v_m \sim \text{seragam sperikal}, \text{ dan } s_{jn} \sim \text{seragam sperikal}$$

symbol N, IG, dan N^+ berturut-turut melambangkan sebaran normal, invers gamma, dan sebaran normal positif.

Selanjutnya sebaran posterior diperoleh melalui perkalian antara prior dengan likelihood. Misalka sebaran untuk $(Y_{ijk} | \theta)$ adalah:

$$Y_{ijk} | \theta \sim N(\mu_{ij}, \sigma^2)$$

dengan

$$\eta_{ij} = \mu + \tau_i + \gamma_j + \sum_{k=1}^m \lambda_k v_{ik} s_{jk}$$

dan θ didefinisikan sebagai

$$\mu, \tau_i, \gamma_j, \lambda_k, v_{ik}, s_{jk}, \sigma^2$$

Sehingga Likelihoodnya diiperoleh sebagai berikut:

$$\begin{aligned} L(\theta) &= \prod_{ijk} (\pi \sigma^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2} (y_{ijk} - \eta_{ij})^2\right\} \\ &= (\pi \sigma^2)^{-abr/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{ijk} (y_{ijk} - \bar{y}_{ij})^2 - \frac{r}{2\sigma^2} \sum_{ij} (\bar{y}_{ij} - \eta_{ij})^2\right\} \end{aligned}$$

Sebaran posterior bersama adalah:

$$\begin{aligned} \pi(\theta | y^n) &\propto L(\theta) \times \pi_{\mu}(\mu) \times \pi_{\tau}(\tau) \times \pi_{\gamma}(\gamma) \times \prod_k \pi_{\lambda_k}(\lambda_k) \times \pi_{v_k}(v_k) \times \pi_{s_k}(s_k) \times \pi_{\sigma^2}(\sigma^2) \\ &= (\pi \sigma^2)^{-abr/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{ijk} (y_{ijk} - \bar{y}_{ij})^2 - \frac{r}{2\sigma^2} \sum_{ij} (\bar{y}_{ij} - \eta_{ij})^2\right\} \\ &\quad \times (\pi \sigma_{\mu}^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_{\mu}^2} (\mu - \mu_{\mu})^2\right\} \times (\pi \sigma_{\tau}^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_{\tau}^2} (\tau_i - \mu_{\tau_i})^2\right\} \\ &\quad \times (\pi \sigma_{\gamma}^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_{\gamma}^2} (\gamma_j - \mu_{\gamma_j})^2\right\} \times (\pi \sigma_{\delta_{ij}}^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_{\delta_{ij}}^2} (\delta_{ij} - \mu_{\delta_{ij}})^2\right\} \\ &\quad \times \frac{\beta^{\alpha}}{\Gamma(\alpha)} \left(\frac{\beta}{\sigma}\right)^{\alpha+1} \exp\left(-\beta/\sigma\right) \end{aligned}$$

Sebaran posterior dari masing-masing parameter diperoleh dari perkalian antara prior dari parameter dengan likelihood.

- Sebaran posterior untuk μ (Liu, 2001)

$$\begin{aligned} \pi(\mu | \theta \text{ lainnya}) &\propto \exp\left\{-\frac{r}{2\sigma^2} \sum_{ij} (\hat{y}_{ij} - \eta_{ij})^2\right\} \times \exp\left\{-\frac{1}{2\sigma_\mu^2} (\mu - \mu_\mu)^2\right\} \\ &\propto \exp\left\{-\frac{r}{2\sigma^2} \sum_{ij} (\hat{y}_{ij} - \mu)^2\right\} \times \exp\left\{-\frac{1}{2\sigma_\mu^2} (\mu - \mu_\mu)^2\right\} \\ &\propto \exp\left\{-\frac{rab\sigma_\mu^2 + \sigma^2}{2\sigma^2\sigma_\mu^2} \left(\mu - \frac{rab\sigma_\mu^2\bar{y}_{\dots} + \sigma^2\mu_\mu}{rab\sigma_\mu^2 + \sigma^2}\right)^2\right\} \end{aligned}$$

$$\mu | \theta \text{ lainnya} \sim N\left(\frac{rab\sigma_\mu^2\bar{y}_{\dots} + \sigma^2\mu_\mu}{rab\sigma_\mu^2 + \sigma^2}, \frac{\sigma^2\sigma_\mu^2}{rab\sigma_\mu^2 + \sigma^2}\right)$$

- Sebaran posterior untuk τ_i

$$\begin{aligned} \pi(\tau_i | \theta \text{ lainnya}) &\propto \exp\left\{-\frac{r}{2\sigma^2} \sum_{ij} (\hat{y}_{ij} - \eta_{ij})^2\right\} \times \exp\left\{-\frac{1}{2\sigma_{\tau_i}^2} (\tau_i - \mu_{\tau_i})^2\right\} \\ &\propto \exp\left\{-\frac{rb\sigma_{\tau_i}^2 + \sigma^2}{2\sigma^2\sigma_{\tau_i}^2} \left(\tau_i - \frac{rb\sigma_{\tau_i}^2\hat{\tau}_i + \sigma^2\mu_{\tau_i}}{rb\sigma_{\tau_i}^2 + \sigma^2}\right)^2\right\} \end{aligned}$$

$$\tau_i | \theta \text{ lainnya} \sim N\left(\frac{rb\sigma_{\tau_i}^2\hat{\tau}_i + \sigma^2\mu_{\tau_i}}{rb\sigma_{\tau_i}^2 + \sigma^2}, \frac{\sigma^2\sigma_{\tau_i}^2}{rb\sigma_{\tau_i}^2 + \sigma^2}\right)$$

- Sebaran posterior untuk γ_j

$$\begin{aligned} \pi(\gamma_j | \theta \text{ lainnya}) &\propto \exp\left\{-\frac{r}{2\sigma^2} \sum_{ij} (\hat{y}_{ij} - \eta_{ij})^2\right\} \times \exp\left\{-\frac{1}{2\sigma_{\gamma_j}^2} (\gamma_j - \mu_{\gamma_j})^2\right\} \\ &\propto \exp\left\{-\frac{ra\sigma_{\gamma_j}^2 + \sigma^2}{2\sigma^2\sigma_{\gamma_j}^2} \left(\gamma_j - \frac{ra\sigma_{\gamma_j}^2\hat{\gamma}_j + \sigma^2\mu_{\gamma_j}}{ra\sigma_{\gamma_j}^2 + \sigma^2}\right)^2\right\} \end{aligned}$$

$$\gamma_j | \theta \text{ lainnya} \sim N\left(\frac{ra\sigma_{\gamma_j}^2\hat{\gamma}_j + \sigma^2\mu_{\gamma_j}}{ra\sigma_{\gamma_j}^2 + \sigma^2}, \frac{\sigma^2\sigma_{\gamma_j}^2}{ra\sigma_{\gamma_j}^2 + \sigma^2}\right)$$

- Sebaran posterior untuk λ_k

$$\pi(\lambda_k | \theta \text{ lainnya}) \propto N\left(\frac{r\sigma_\lambda^2 \sum_{ij} v_{ik} s_{jk} \bar{y}_{ij} + \sigma^2 \mu_\lambda}{r\sigma_\lambda^2 + \sigma^2}, \frac{\sigma_\lambda^2 \mu_\lambda}{r\sigma_\lambda^2 + \sigma^2}\right)$$

Untuk $\lambda_{k-1} \geq \lambda_k \geq \lambda_{k+1}$ dan diasumsikan $\lambda_0 = \infty$ dan $\lambda_{m+1} = 0$.

- Sebaran posterior untuk v_k dengan $k=1, 2, \dots, m$

$$\begin{aligned} \pi(v_k | \theta \text{ lainnya}) &\propto \exp\left\{\frac{r\lambda_k}{\sigma^2} \left(\sum_i \sum_j v_{ik} s_{jk} \bar{y}_{ij}\right)\right\} \\ &\propto \exp\left\{\frac{r\lambda_k}{\sigma^2} v_k^T \dot{v}_k\right\} \end{aligned}$$

Dengan $\dot{v}_k = \sum_j s_{jk} \bar{y}_{ij}$.

Untuk model AMMI, karena v_n harus orthogonal terhadap vector $\mathbf{1}_a$ dan v yang lain, $A-v_k$, ada matrik \mathbf{H}_k berukuran $a \times (a-m)$ dimana kolom dari \mathbf{H}_k adalah suatu gugus vektor ortonormal dan orthogonal terhadap $\mathbf{1}_a$ dan $A-v_k$. jika didefinisikan $v_k^* = \mathbf{H}_k^T v_k$ yang merupakan transformasi linier satu-satu, sebaran posterior dari v_k^* dengan $v_k^* \in V_{a-m}$ adalah:

$$\begin{aligned} \pi(v_k^* | \theta \text{ lainnya}) &\propto \exp\left\{\frac{r\lambda_k}{\sigma^2} v_k^{*T} \mathbf{H}_k \mathbf{H}_k^T \dot{v}_k\right\} \\ &\propto \exp\left\{\frac{rc_k \lambda_k}{\sigma^2} v_k^{*T} \tilde{v}_k\right\} \end{aligned}$$

dengan $\tilde{v}_k = c_k^{-1} \mathbf{H}_k^T \dot{v}_k$ dan $c_k = \sqrt{\dot{v}_k^T \mathbf{H}_k \mathbf{H}_k^T \dot{v}_k}$.

Selanjutnya akan diperoleh $v_k^* | \theta \text{ lainnya} \sim FM(a-m, \frac{rc_k \lambda_k}{\sigma^2}, \tilde{v}_k)$, dengan FM adalah sebaran von Mises Fisher.

- Sebaran posterior untuk s_k dengan $k=1, 2, \dots, m$

$$\begin{aligned} \pi(s_k | \theta \text{ lainnya}) &\propto \exp\left\{\frac{r\lambda_k}{\sigma^2} \left(\sum_i \sum_j v_{ik} s_{jk} \bar{y}_{ij}\right)\right\} \\ &\propto \exp\left\{\frac{r\lambda_k}{\sigma^2} s_k^T \dot{s}_k\right\} \end{aligned}$$

Dengan $\dot{s}_k = \sum_i v_{ik} \bar{y}_{ij}$.

Dengan cara yang hampir sama seperti dalam menentukan sebaran posterior untuk v_k , sebaran dari $s_k^* | \theta \text{ lainnya} \sim FM(b-m, \frac{rd_k \lambda_k}{\sigma^2}, \tilde{s}_k)$,

dimana $\tilde{s}_k = d_k^{-1} \mathbf{R}_k^T \dot{s}_k$ dan $d_k = \sqrt{\dot{s}_k^T \mathbf{R}_k \mathbf{R}_k^T \dot{s}_k}$ serta \mathbf{R}_k berukuran $b \times (b-$

m) dimana kolom dari R_k adalah suatu gugus vector ortonormal dan orthogonal terhadap $\mathbf{1}_b$ dan S - s_k .

- Sebaran posterior untuk σ^2

$$\begin{aligned} \pi(\sigma^2 | \theta \text{ lainnya}) &\propto L(\theta) \times \pi_{\sigma^2}(\sigma^2 | \alpha_\sigma, \beta_\sigma) \\ &\propto (\pi_{\sigma^2})^{abr/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{ijk} (y_{ijk} - \eta_{ij})^2\right\} \times (\sigma^2)^{-(\alpha_\sigma+1)} \exp\left\{-\frac{\beta_\sigma}{\sigma^2}\right\} \\ &\propto (\sigma^2)^{-(abr/2+\alpha_\sigma+1)} \exp\left\{-\frac{1}{\sigma^2} \left(\beta_\sigma + \frac{1}{2} \sum_{ijk} (y_{ijk} - \eta_{ij})^2\right)\right\} \end{aligned}$$

Untuk menduga parameter model AMMI melalui pendekatan Bayesian selanjutnya dihitung dengan *Gibbs sampling*. Proses yang dilakukan pada tahap ini adalah:

1. Dihitung nilai dugaan awal dari parameter model menggunakan metode kuadrat terkecil (MKT):

- a. $\hat{\mu} = \frac{1}{abr} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r y_{ijk}$

- b. $\hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..}$

- c. $\hat{\gamma}_j = \bar{y}_{.j} - \bar{y}_{..}$

- d. $\hat{\delta}_{ij} = \bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$

- e. Penguraian nilai singular dari $\hat{\delta}_{ij} = VLS^T$, $\hat{\lambda}_k$ merupakan unsur diagonal dari L, \hat{v}_k merupakan vektor dari kolom matriks V, dan \hat{s}_k merupakan vektor dari kolom matriks S.

2. Dibangkitkan sebaran posterior untuk $\theta = (\hat{\mu}, \hat{\tau}_i, \hat{\gamma}_j, \hat{\lambda}_k, \hat{v}_k, \hat{s}_k)$ menggunakan *Gibbs sampling*. Nilai awal dari setiap parameter menggunakan nilai dugaan awal yang diperoleh seperti pada tahap 1.
3. Nilai rata-rata dari sebaran posterior yang dibangkitkan pada tahap 2 digunakan untuk menduga μ, τ, γ , dan λ , dimana:

- a. $\tilde{\mu} = \frac{1}{m} \sum_{l=1}^m \mu^{(l)}$

- b. $\tilde{\tau}_i = \frac{1}{m} \sum_{l=1}^m \tau_i^{(l)}$

c. $\tilde{\gamma}_j = \frac{1}{m} \sum_{l=1}^m \gamma_{jl}$

d. $\tilde{\lambda}_k = \frac{1}{m} \sum_{l=1}^m \lambda_{kl}$, untuk $k=1, 2, \dots, m$

Untuk menduga parameter v dan s dilakukan dengan tahapan sebagai berikut:

1. Dibuat matriks B yang berukuran $p \times q$ dengan $p \geq q$ dimana kolom dari B dibentuk dari vektor v_k atau s_k .
2. Dihitung $\bar{B} = \frac{1}{m} \sum_{l=1}^m B_l$.
3. Dilakukan penguraian nilai singular untuk matriks \bar{B} sehingga akan diperoleh $\bar{B} = LDR^T$.

$\hat{B} = LR^T$ adalah matriks yang unsur kolomnya merupakan dugaan dari parameter v_k atau s_k .

Banyaknya KUI yang akan dipertahankan dalam model AMMI dapat ditentukan menggunakan kriteria BIC (*Bayesian Information Criteria*) (Liu, 2001). Banyaknya KUI yang akan digunakan yakni model yang memiliki nilai BIC paling kecil.

Pada bagian selanjutnya disajikan artikel tentang penggunaan pendekatan bayesian untuk pendugaan matriks Interaksi pada model AMMI.

Halaman ini sengaja dikosongkan

Pendekatan Metode Bayes untuk Pendugaan Pengaruh Interaksi pada Model AMMI

Pika Silvianti¹, Khairil A. Notodiputro², dan I Made Sumertajaya²

Abstract

Multi-locations trials play an important role in plant breeding and agronomic research. Study concerning genotype-environment interaction needed in selection of genotype to be released. AMMI (Additive Main effect and Multiplicative Interaction) is one of statistical technique to analyze data from multi-locations trials. The analysis of AMMI is a combining analysis between additive main effect and principal component analysis. Multi-location sampling data which were collected several years on several planting season used to be analyzed separately. To obtain more comprehensive information of multi-location sampling data, an analysis which combines all the information in several years is needed. One of the alternatives is the Bayesian approach. This method utilizes initial information on the estimated parameters and information from samples. The simulation states that prediction with Bayesian methods will produce a better estimator, because MSE of the Bayesian estimator is smaller the MSE estimator generated using least squares method.

Keywords: AMMI, Bayes

PENDAHULUAN

Latar Belakang

Percobaan di multi-lokasi merupakan teknik percobaan yang sering dilakukan dan sangat penting dalam bidang pemuliaan tanaman. Percobaan semacam ini melibatkan dua faktor utama yaitu genotipe tanaman dan kondisi lingkungan (lingkungan: tempat (*site*), musim, perlakuan agronomis (*agronomy treatment*)). Metode statistika yang biasa digunakan untuk analisis kestabilan terhadap hasil percobaan multilokasi adalah AMMI (*Additive Main Effect and Multiplicative Interaction*). Data percobaan multilokasi ini dikumpulkan dari beberapa tahun di beberapa musim tanam. Namun, analisis dari data percobaan multilokasi ini masih dilakukan secara terpisah antara data tahun satu dengan tahun yang lainnya. Agar informasi dari data percobaan multilokasi dapat diperoleh secara

¹ Mahasiswa S2 Program Studi Statistika SPs IPB

² Departemen Statistika, FMIPA IPB

lebih menyeluruh, maka perlu adanya suatu analisis yang menggabungkan informasi-informasi dalam beberapa tahun tersebut. Salah satu alternatif analisis yang dapat kita gunakan adalah pendekatan Bayes. Metode ini memanfaatkan informasi awal tentang parameter yang akan diduga dan informasi dari contoh.

Tujuan

Beberapa tujuan dari artikel ini antara lain:

1. Mempelajari kinerja dari dugaan parameter yang dihasilkan dengan metode Bayes.
2. Menentukan genotipe stabil berdasarkan dugaan metode Bayes.

METODOLOGI

Data

Data yang akan digunakan dalam penelitian ini ada dua jenis, data pertama adalah data yang dibangkitkan dalam program simulasi. Data simulasi dibangun dari model percobaan multilokasi dengan ragam contoh di setiap lokasi diasumsikan sama. Parameter yang dibutuhkan untuk membangkitkan data dalam simulasi ini adalah nilai tengah hasil produksi, pengaruh faktor genotipe, keragaman lokasi percobaan kecil ($\sigma_{\gamma_j}^2 = 1$) dan keragaman lokasi percobaan sedang ($\sigma_{\gamma_j}^2 = 5$), keragaman interaksi kecil ($\sigma_{\delta_{ij}}^2 = 1$) dan keragaman interaksi sedang ($\sigma_{\delta_{ij}}^2 = 5$), serta keragaman galat ($\sigma_{\varepsilon}^2 = 1$). Faktor genotipe diasumsikan tetap, sesuai dengan kondisi pada data riil. Dalam simulasi ditentukan jumlah lokasi percobaan sebanyak 20, dibuat simulasi 100 set data dari model di atas.

Data kedua adalah data riil yang digunakan untuk penerapan yang merupakan data dari percobaan internasional untuk gandum yang dilakukan oleh program CIMMYT (*International Maize and Wheat Improvement Center*) pada 12 genotipe yang ditanam di empat lokasi dengan 4 blok pada dua tahun berturut-turut yaitu tahun 2005 dan tahun 2006 dan data hasil penelitian oleh Konsorsium padi Nasional, yaitu Penelitian Interaksi antara Genotipe dengan Lingkungan pada galur harapan padi sawah tahun 2008 menggunakan 14 galur padi dimana 11 galur (1 galur berasal dari BATAN, 5 galur dari BB Padi, 1 galur dari Biogen, dan 4 galur dari IPB), dengan 3 varietas pembanding (Gilirang, INPARI1, dan Ciherang) yang ditanam pada 21 lokasi.

Metode Pendugaan Parameter

Pendugaan parameter dapat dilakukan dengan menggunakan algoritma *Gibbs sampling*. *Gibbs sampling* adalah suatu teknik untuk membangkitkan peubah acak dari sebaran (marjinal) secara tidak langsung, tanpa perlu menghitung fungsi kepekatannya (Casella & George, 1992). Dengan menggunakan teknik *Gibbs sampling*, kita dapat menghindari perhitungan yang sulit. *Gibbs sampling* merupakan salah satu metode untuk membangun algoritma *Markov Chain Monte Carlo* (MCMC). Algoritma MCMC diimplementasikan dengan cara mengambil contoh berulang-ulang dari p sebaran posterior bersyarat $[\vartheta_1/\vartheta_2, \dots, \vartheta_p], \dots, [\vartheta_p/\vartheta_1, \dots, \vartheta_{p-1}]$ (Albert, 2007).

Nilai awal yang digunakan adalah nilai dugaan pengaruh interaksi dengan menggunakan metode kuadrat terkecil. Misalkan θ^l untuk $l=1, \dots, m$ adalah contoh yang dibangkitkan dengan *Gibbs sampling* untuk model percobaan multilokasi. Rataan dari contoh digunakan untuk menduga μ, τ, γ , dan δ (Liu, 2001).

$$\tilde{\mu} = \frac{1}{m} \sum_{l=1}^m \mu^l; \tilde{\tau}_i = \frac{1}{m} \sum_{l=1}^m \tau_i^l; \tilde{\gamma}_j = \frac{1}{m} \sum_{l=1}^m \gamma_j^l; \tilde{\delta}_{ij} = \frac{1}{m} \sum_{l=1}^m \delta_{ij}^l$$

Kriteria Evaluasi

Nilai dugaan terhadap pengaruh interaksi dievaluasi menggunakan dua kriteria, yaitu:

1. $Bias(\hat{\delta}) = E(\hat{\delta}) - \delta$
2. $MSE = E\left((\hat{\delta} - \delta)^2\right) = var(\hat{\delta}) + Bias^2(\hat{\delta})$

Setelah nilai Bias dan MSE dari kedua metode didapatkan, maka akan dilakukan perbandingan terhadap nilai bias dan MSE.

Simulasi

Kinerja dari penduga Bayes untuk pengaruh interaksi dievaluasi dengan melakukan simulasi. Algoritma *gibbs sampling* dilakukan sebanyak $l=1000$ dan 5000 untuk membangkitkan sebaran posterior dari masing-masing parameter dengan periode *burn-in* sebanyak 100 dan 1000 .

Tahapan simulasi:

1. Tetapkan nilai-nilai parameter berikut : $\mu, \sigma_\gamma^2, \sigma_{\delta_{ij}}^2, \sigma_\varepsilon^2, \alpha_\sigma, \beta_\sigma$
2. Bangkitkan $\tau_i, \gamma_j, \varepsilon_{ijk}$, dan δ_{ij}
3. Dapatkan nilai Y berdasarkan model $y_{ijk} = \mu + \tau_i + \gamma_j + \delta_{ij} + \varepsilon_{ijk}$

4. Hitung nilai dugaan parameter dengan metode MKT ($\hat{\mu}, \hat{\tau}_i, \hat{\gamma}_j, \hat{\delta}_{ij}, \hat{\sigma}^2$) dan ragam dari masing-masing parameter model $\mu^2, \sigma_{\tau_i}^2, \sigma_{\gamma_j}^2, \sigma_{\delta_{ij}}^2$ gunakan sebagai nilai awal untuk masuk ke algoritma *gibbs sampling*
5. Hitung dugaan parameter model dengan metode Bayes menggunakan algoritma *gibbs sampling*

i. Tentukan nilai awal $\theta^0 = (\mu^{(0)}, \tau_i^{(0)}, \gamma_j^{(0)}, \delta_{ij}^{(0)}, \sigma^{2(0)})$

ii. Ulangi langkah untuk $l = 1, 2, \dots, 1000$

a) Bangkitkan μ^C dari $\pi(\mu | \tau_i^{(l-1)}, \gamma_j^{(l-1)}, \delta_{ij}^{(l-1)}, \sigma^{2(l-1)})$

$$\pi(\mu | \tau_i, \gamma_j, \delta_{ij}, \sigma^2) \propto \exp\left\{-\frac{rab\sigma_{\mu}^2 + \sigma^2}{2\sigma^2\sigma_{\mu}^2} \left(\mu - \frac{rab\sigma_{\mu}^2\bar{y}_{...} + \sigma^2\mu_{\mu}}{rab\sigma_{\mu}^2 + \sigma^2}\right)^2\right\}$$

b) Bangkitkan σ^{2C} dari $\pi(\sigma^2 | \mu^{(l-1)}, \tau_i^{(l-1)}, \gamma_j^{(l-1)}, \delta_{ij}^{(l-1)})$

$$\pi(\sigma^2 | \mu, \tau_i, \gamma_j, \delta_{ij}) \propto IG\left[\frac{abr}{2} + \alpha_{\sigma}, \beta_{\sigma} + \frac{1}{2} \sum_{ijk} (\mu_{ijk} - \eta_{ij})^2\right]$$

c) Bangkitkan τ_i^C dari $\pi(\tau_i | \mu^l, \gamma_j^{(l-1)}, \delta_{ij}^{(l-1)}, \sigma^{2(l-1)})$

$$\pi(\tau_i | \mu, \gamma_j, \delta_{ij}, \sigma^2) \propto \exp\left\{-\frac{rb\sigma_{\tau_i}^2 + \sigma^2}{2\sigma^2\sigma_{\tau_i}^2} \left(\tau_i - \frac{rb\sigma_{\tau_i}^2\hat{\tau}_i + \sigma^2\mu_{\tau_i}}{rb\sigma_{\tau_i}^2 + \sigma^2}\right)^2\right\}$$

d) Bangkitkan γ_j^C dari $\pi(\gamma_j | \mu^{(l)}, \tau_i^{(l)}, \delta_{ij}^{(l-1)}, \sigma^{2(l-1)})$

$$\pi(\gamma_j | \mu, \tau_i, \delta_{ij}, \sigma^2) \propto \exp\left\{-\frac{ra\sigma_{\gamma_j}^2 + \sigma^2}{2\sigma^2\sigma_{\gamma_j}^2} \left(\gamma_j - \frac{ra\sigma_{\gamma_j}^2\hat{\gamma}_j + \sigma^2\mu_{\gamma_j}}{ra\sigma_{\gamma_j}^2 + \sigma^2}\right)^2\right\}$$

e) Bangkitkan δ_{ij}^C dari $\pi(\delta_{ij} | \mu^{(l)}, \tau_i^{(l)}, \gamma_j^{(l)}, \sigma^{2(l-1)})$

$$\pi(\delta_{ij} | \mu, \tau_i, \gamma_j, \sigma^2) \propto \exp\left\{-\frac{r\sigma_{\delta_{ij}}^2 + \sigma^2}{2\sigma^2\sigma_{\delta_{ij}}^2} \left(\delta_{ij} - \frac{r\sigma_{\delta_{ij}}^2\hat{\delta}_{ij} + \sigma^2\mu_{\delta_{ij}}}{r\sigma_{\delta_{ij}}^2 + \sigma^2}\right)^2\right\}$$

iii. Kembalikan $\theta^l = (\mu^{(l)}, \tau_i^{(l)}, \gamma_j^{(l)}, \delta_{ij}^{(l)})$

6. Hitung nilai rata-rata dari masing-masing sebaran posterior, gunakan sebagai penduga parameter model multilokasi $(\tilde{\mu}, \tilde{\tau}_i, \tilde{\gamma}_j, \tilde{\delta}_{ij})$
7. Evaluasi keakuratan penduga interaksi dengan mengukur besarnya bias
8. Evaluasi presisi penduga interaksi dengan mengukur besarnya MSE

Penerapan

Data riil dari CIMMYT dan BB Tanaman Padi Sukamandi digunakan untuk menerapkan metode Bayes dalam pendugaan parameter model AMMI. Tahapannya sebagai berikut:

1. Mencari informasi prior
2. Data Tahun Kedua digunakan untuk analisis AMMI dan mengevaluasi kestabilan genotipe
 - a. Duga parameter model AMMI ($\hat{\mu}, \hat{\tau}_i, \hat{\gamma}_j, \hat{\delta}_{ij}$) serta ragam (σ^2) dengan MKT
 - b. Gunakan dugaan MKT sebagai nilai awal untuk menghitung dugaan parameter dengan metode Bayes $(\tilde{\mu}, \tilde{\tau}_i, \tilde{\gamma}_j, \tilde{\delta}_{ij})$
 - c. Susun Matriks interaksi, gunakan matriks interaksi untuk analisis AMMI
 - d. Tentukan genotipe stabil dan genotipe spesifik berdasarkan metode AMMI

HASIL DAN PEMBAHASAN

Pada Tabel 1 disajikan rata-rata keseluruhan bias dan MSE dari penduga pengaruh interaksi menggunakan MKT dan Bayes. Bias dari penduga Bayes dan penduga MKT memiliki nilai yang bervariasi.

Namun secara umum dapat kita lihat, nilai absolut bias dari penduga Bayes relatif lebih kecil dibandingkan dengan bias penduga MKT. Hal yang sama terjadi pada MSE, dimana pada berbagai kondisi ragam lokasi dan ragam interaksi MSE dari penduga Bayes nilainya selalu lebih kecil dari MSE penduga MKT yang merupakan indikasi bahwa metode Bayes memiliki performa lebih baik dibandingkan metode MKT karena tingkat kesalahan yang dihasilkan oleh metode Bayes relatif lebih kecil.

Terlihat bahwa untuk ragam lokasi yang sama, kemampuan metode Bayes memperbaiki kinerja dugaannya cenderung menurun dengan meningkatnya nilai ragam interaksi. Sedangkan pada nilai ragam interaksi yang sama, kemampuan metode Bayes memperbaiki dugaan cenderung meningkat dengan semakin besarnya ragam lokasi.

Simulasi juga dilakukan untuk mengevaluasi kinerja metode Bayes dalam mengklasifikasikan genotipe-genotipe stabil dengan menggunakan Biplot AMMI. Karena proses membuat Biplot AMMI membutuhkan tahapan yang sangat panjang, untuk itu simulasi ini tidak dilakukan sebanyak simulasi dalam

pendugaan parameter. Simulasi penentuan klasifikasi genotipe menggunakan Biplot AMMI dilakukan pada kondisi keragaman lokasi kecil ($\sigma_{\gamma_j}^2 = 1$) dan keragaman interaksi sedang ($\sigma_{\delta_{ij}}^2 = 5$), serta pada kondisi keragaman lokasi besar ($\sigma_{\gamma_j}^2 = 5$) dan keragaman interaksi kecil ($\sigma_{\delta_{ij}}^2 = 1$). Kondisi ini dipilih karena adanya perbaikan yang cukup ekstrim dari dugaan metode Bayes yang diberikan pada kedua kondisi ini sebagaimana dijelaskan pada Tabel 1.

Tabel 1. Rata-Rata Bias dan MSE pada masing-masing kondisi simulasi

Ragam Lokasi	Ragam Interaksi	N	Burn-in	Bias		MSE		Persentase Perbaikan dugaan (%)*
				Bayes	MKT	Bayes	MKT	
				$\tilde{\delta}_{ij}$	$\hat{\delta}_{ij}$	$\tilde{\delta}_{ij}$	$\hat{\delta}_{ij}$	
1	1	1000	100	0.0013	-0.0016	1.1499	2.1186	45.72
		5000	1000	-0.0026	-0.0054	1.0931	2.1107	48.21
1	5	1000	100	0.0027	0.0230	7.8567	8.8756	11.48
		5000	1000	-0.0122	-0.0487	8.7886	9.1179	3.61
5	1	1000	100	-0.0025	-0.0032	1.2428	13.4729	90.78
		5000	1000	0.0014	0.0015	1.2505	13.3637	90.64
5	5	1000	100	0.0010	0.0023	7.5335	20.5016	63.25
		5000	1000	0.0294	-0.0444	8.1250	20.6310	60.62

*: Persentase Perbaikan dugaan = $\frac{MSE \hat{\delta}_{ij} - MSE \tilde{\delta}_{ij}}{MSE \hat{\delta}_{ij}} \times 100\%$

Tabel 2. Simulasi untuk Klasifikasi Genotipe dengan Biplot AMMI

Ragam Lokasi	Ragam Interaksi	Genotipe Stabil		
		Parameter	MKT	Bayes
		δ_{ij}	$\hat{\delta}_{ij}$	$\tilde{\delta}_{ij}$
1	5	G13	G7,G13	G13
		G2, G9	G9,G5	G2,G9,G7
5	1	G11,G13	G11,G13	G11,G13
		G10	-	G10

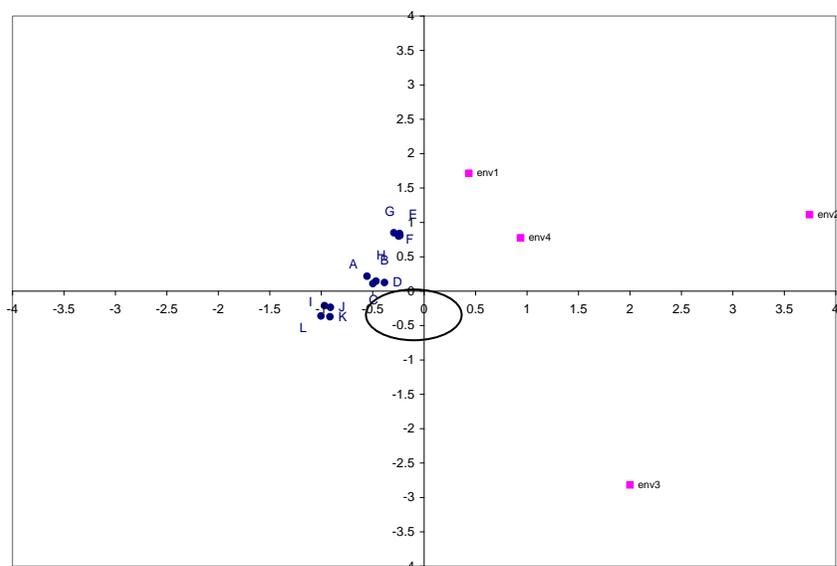
Pada Tabel 2, disajikan hasil simulasi klasifikasi genotipe menggunakan Biplot AMMI. Terlihat bahwa genotipe-genotipe yang diklasifikasikan stabil oleh metode Bayes, tidak terlalu berbeda dengan genotipe yang stabil dalam kondisi sesungguhnya (parameter). Sedangkan pada klasifikasi menggunakan MKT ada

Model AMMI Kini dan Yang Akan Datang

beberapa genotipe yang digolongkan stabil, namun pada keadaan sesungguhnya tidak stabil. Hal ini juga terjadi di beberapa klasifikasi genotipe dengan metode Bayes, namun kesalahan klasifikasinya lebih sering terjadi pada MKT.

Penerapan Data Gandum

Data yang digunakan untuk ilustrasi berikut merupakan data percobaan internasional untuk gandum yang dilakukan oleh program CIMMYT (*International Maize and Wheat Improvement Center*). Pada Gambar 1 berikut disajikan Biplot AMMI dengan matriks pengaruh interaksi menggunakan pendugaan dengan pendekatan Bayes.



Gambar 1. Biplot AMMI Data Gandum

Perhitungan selang kepercayaan normal ganda pada taraf $\alpha = 0.05$ menghasilkan ellips dengan jari-jari panjang 0.47 dan jari-jari pendek 0.38. Terlihat bahwa genotipe D masuk ke dalam daerah kepercayaan ellips, yang berarti genotipe ini dinyatakan sebagai genotipe stabil di semua lokasi percobaan. Sedangkan genotipe A,B,C,E,F,G,H,I,J,K,L merupakan genotipe yang tidak stabil karena posisinya berada di luar daerah kepercayaan ellips. Hasil biplot AMMI menggunakan penduga dengan metode MKT memberikan kesimpulan dimana tidak ada genotipe yang dikategorikan stabil.

Penerapan Data Padi

diduga dengan metode Bayes relatif tidak berbeda dengan kondisi data aslinya (parameter model) dibandingkan dengan klasifikasi genotipe yang dihasilkan oleh dugaan MKT. Untuk itu dalam menentukan genotipe yang stabil didalam suatu percobaan, keberadaan informasi prior perlu dipertimbangkan dalam analisis.

UCAPAN TERIMAKASIH

Tulisan ini bagian dari penelitian HIBAH PENELITIAN PASCASARJANA, No. 41/I3.24.4/SPK/BG-PD/2009, tanggal 30 Maret 2009.

PUSTAKA

- Albert, Jim.** 2007. *Bayesian Computation with R*.
<http://www.springerlink.com/content/t43r812716455567/> [3 Juni 2009].
- Casella, G., and E. I. George.** 1992. *Explaining the Gibbs sampler*. *American Statistikian*. 46, 167-174. <http://www.jstor.org/sTabel/2685208?origin=JSTOR-pdf> [29 Mei 2009].
- Berger, J.O.** 1985. *Statistikal Decision Theory and Bayesian Analysis, 2nd edition*. Springer Verlag (New York).
- Cotes, Jose M., J. Crossa, A. Sanches, and P.L. Cornelius.** 2006. *A Bayesian Approach for Assessing the Stability of Genotypes*. <http://crop.scijournals.org/cgi/content/full/46/6/2654> [2 Juni 2008]
- Liu, Genzhou.** 2001. *Bayesian Computation for Linear-Bilinear Model*. University of Kentucky. Disertasi.

PENUTUP

11

PENUTUP

Fokus utama penelitian ini adalah pengembangan utama model selesai adaptabilitas melalui model AMMI. Hasil penelitian ini secara menyeluruh dapat kita pandang dalam empat elemen yang saling terkait. Keempat elemen tersebut saling terait satu sama lain yaitu (1) Penanganan ketersediaan atau kelengkapan data (2) Penanganan permasalahan Mode AMMI pada Distribusi Normal dan Pengembangan Inferensi pada Model AMMI, (3) Pengembangan/perluasan Model AMMI (*more general on AMMI Models*) dan (4) mengenai pengembangan komputasi dan aplikasinya.

11.1 Ringkasan

Keseluruhan hasil pengembangan model AMMI dapat digambarkan oleh keempat hal tersebut. Secara ringkas keempat elemen tersebut diuraikan sebagai berikut:

- (1) **Ketersediaan data.** Pengembangan ini meliputi penanganan masalah ketidaklengkapan data dan pengamatan hilang dan penggabungan respon untuk pengamatan daya hasil. Data hilang dapat diselesaikan dengan **EM-AMMI dan Metode g-connected**, sedangkan penggabungan respon dapat dilakukan dengan membuat **Indeks Performa Tanaman (IPT)**
- (2) **Penanganan permasalahan Mode AMMI pada Distribusi Normal dan Pengembangan Inferensi pada Model AMMI.** Pengembangan pada model acak dilakukan antara lain (1) **Pengembangan model AMMI campuran** bagi percobaan multilokasi adalah (i) pedugaan pengaruh interaksi menggunakan BLUP dan (ii) dekomposisi matriks interaksi menggunakan model Analisis Faktor. Jika lingkungan yang dilibatkan dalam percobaan merupakan faktor acak, maka secara teori model campuran lebih tepat digunakan. Sedangkan berdasarkan struktur model dan kemampuan klasifikasi serta kemampuan pendugaan, maka model campuran lebih baik dibandingkan model tetap karena menghasilkan struktur model yang lebih sederhana dengan kemampuan pendugaan dan klasifikasi yang sama. (2) **Pemodelan IGL dengan menggunakan Structural Equation Modeling (SEM).** Melalui Struktural Equation Modeling dapat dilakukan analisis GEI untuk daya hasil dengan memasukkan komponen daya hasil dan kovariat genotypic \times lokasi yang berguna untuk mengindenfikasi faktor-faktor yang menyebabkan nyataanya efek interaksi genotip \times lokasi dalam analisis Ragam AMMI. Bila terdapat masalah tidak terpenuhi asumsi pada SEM AMMI maka metode *Partial Least Square* dapat menjadi solusi. Tterkait dengan pengembangan inferensia pada model AMMI adalah (3) **Pengujian sub-hipotesis pada IGL** berguna untuk mengetahui genotip yang berkontribusi terhadap pengaruh interaksi genotip dan lingkungan. (4) **Inferensi Kestabilan Genotipe pada Model AMMI** menggunakan metode Resampling Bootstrap. Melalui resampling akan memberikan inferensia untuk interpretasi hasil biplot AMMI. Metode resampling bootstrap dapat diperoleh daerah kepercayaan skor KUI_1^* dan KUI_2^* setiap genotipe. Analisis AMMI menggunakan resampling bootstrap cocok untuk mendeteksi kestabilan pada percobaan multilokasi apabila ragam genotipe pada biplot kecil.

(3) Pengembangan/peluasannya Model AMMI (more general on AMMI Models).

Pengembangan perluasan ini terkait pada respon model AMMI baik itu mengantisipasi adanya pencilan, tidak terpenuhinya asumsi sebaran normal atau karena data pengamatan dalam bentuk persentase, cacahan, atau pada kelas distribusi/sebaran lain dalam keluarga eksponensial. (1) **Pengembangan Model Kekas Pencilan.** Pendekatan *robust alternating regression* pada model robust-AMMI mampu menangani pencilan dan sekaligus identifikasi adanya pencilan didalam data tabel dua arah atau pun pada matriks interaksi yang akan didekomposisi. (2) **Model Kestabilan Non-Parametrik** digunakan untuk mengatasi masalah asumsi galat dan pengaruh interaksi yang menyebar normal. Untuk itu kita dapat dipilih pendekatan Non-Parametrik, antara lain indeks stabilitas *theranasu*. (3) **Perluasan Model AMMI Terampai.** Untuk data pengamatan yang tidak berdistribusi normal bukan karena adanya pencilan atau hal lain, tetapi karena pencatatan dalam bentuk persen atau cacahan dapat dipilih metode *Generalized AMMI* (G-AMMI). Metode ini memberikan informasi menarik terutama pada sebaran Poisson yaitu tentang Odds Ratio, terutama untuk data/pengamatan serangan hama atau penyakit. (4) **Pengembangan model Bayesian AMMI.** Topik ini merupakan keberlanjutan dari upaya mengatasi ketahomogenan ragam sekaligus upaya mendapatkan inferensi atas hasil AMMI. Didalamnya juga terkait penggunaan distribusi yang unik bagi pengaruh interaksi dan konsep resampling dalam komputasi Bayesian.

(4) Penyediaan Aplikasi Komputer.

Yang tidak kalah penting adalah mengintegrasikan pengembangan metode metode analisis interaksi genotipe interaksi untuk seleksi adaptasi secara komprehensif. Integrasi ini diperlukan mengingat genetika kuantitatif telah sukses membantu pemulia tanaman, tetapi aplikasi statistika sering sulit dimengerti oleh seseorang yang bekerja pada genetika molekular. Ini merupakan kegiatan pengembangan program aplikasi ini ditujukan untuk mengintegrasikan beberapa metode analisis interaksi genotipe interaksi secara komprehensif. Penyediaan aplikasi komputer sebagai upaya untuk menyediakan software alternatif dengan beberapa keunggulan antara lain (i) dikembangkan menggunakan software R yang gratis, (ii) dilengkapi Graphical User Interface (GUI) untuk memudahkan penggunaan, (iii) memiliki kekhususan pada pengembangan metode AMMI. Terakhir, bersama buku ini dilampirkan hasil pengembangan komputasi yang bernama IPB AMMI-R. Penghargaan setinggi-tingginya kami sampaikan pada Achi Rinaldi dan Agus M. Sholeh yang telah mengembangkannya.

11.2 Pengembangan Lebih Lanjut (*further works*)

Pengembangan penelitian dapat terjadi pada beberapa hal. Pertama, selain masalah-masalah pemuliaan dan proteksi tanaman, masalah lain yang melibatkan prevalensi dapat didekati dengan model bilinear ini. Beberapa masalah pembagunan seperti masalah sosial-kependudukan dan ekonomi tentu melibatkan interaksi-interaksi yang tidak mudah dijelaskan. Perluasan pertama ini berkenaan dengan perluasan pada bidang penerapan model. Tidak hanya pada bidang pertanian dan pemuliaan, mungkin pada bidang-bidang sosial, politik dan pemasaran. Perluasan kedua dapat dilakukan secara khusus pada model GAMMI untuk data cacahan, sebaran Poisson menemui fenomena menarik dengan kasus-kasus seperti *Zero Inflated Poisson* ataupun *Negative Binomial Counts*. Hal ini membawa kita pada suatu kelas pemodelan statistika yang lebih luas yaitu interaksi/asosiasi tabel dua arah baris-kolom (*Row-Column Two Ways Tabel Assosiasion*) yang bersisikan data cacahan.

Selain itu, hasil tentang Circular Statistics, sebuah ide lama yang inspiratif baik dalam penelitian ini ataupun untuk penekjaan penelitian berikutnya. Pemanfaatan untuk memberikan inferensia pada sebuah titik pada biplot pada AMMI melalui distribusi Sirkular Normal, telah menunjukkan bahwa perlu cara pandang yang sedikit berbeda terhadap besaran dan arah pada suatu pengukuran. Suatu kajian matematis dan sedikit filofosis menarik untuk dilakukan terkait dengan cara pandang yang berbeda terhadap konsep besaran dan arah pada suatu pengukuran menurut *Circular Statistics*. Tidak hanya itu *Circular Statistics* penting dipahami dalam pengembangan komputasi Bayesian AMMI terutama terkait beberapa sebaran akarciri dan vektor ciri.

Disamping pengembangan beberapa kajian perbandingan dapat dilakukan lebih lanjut, antara lain, seperti (1) perbandingan performa BLUP dan BAYES pada pendugaan matriks interaksi, dan karena penggunaan Bayes yang telah dilakukan baru pada kondisi keragaman data yang sangat terbatas, maka perlu dilakukan (2) pengembangan metode Bayes pada kondisi keragaman data yang lebih umum. Disamping itu perlu pula pengembangan Bayesian pada pemodelan AMMI secara lengkap dalam mendekomposisi interaksi bukan hanya pada pendugaan matriks interaksi. Topik-topik ini potensial untuk dijadikan topik riset dasar.

Atas manfaat buku ini dan pengembangan lebih lanjut menjadi harapan besar kami bagi kemajuan ilmu pengetahuan, statistika, matematika, komputasi dan pembangunan pertanian Indonesia. Semoga.

DAFTAR PUSTAKA

- Albert, Jim. 2007. Bayesian Computation with R. [http://www.springerlink.com / content/t43r812716455567/](http://www.springerlink.com/content/t43r812716455567/).
- Alberts, Martin J. A. 2004. A Comparison Of Statistical Methods to Describe Genotype X Environment Interaction and Yield Stability in Multi-Location Maize Trials. Bloemfontein: University Of The Free State. <http://etd.uovs.ac.za/ETD-db//theses/available/etd-09072005-084932/unrestricted/ALBERTSMJA.pdf> .
- Berger, J.O. 1985. Statistical Decision Theory and Bayesian Analysis, 2nd edision. Springer Verlag. NY.
- Casella, G., and E. I. George. 1992. Explaining the Gibbs sampler. American Statistician. 46, 167-174. <http://www.jstor.org/stable/2685208?origin=JSTOR-pdf> [29 Mei 2009].
- Chernick MR. 2008. Bootstrap Methods: a Guide for Practitioners and Researches, 2nd ed. New Jersey: John Wiley & Son, Inc.
- Cornelius PL, Crossa J, Seyedsadr MS. 1996, Statistical Test and Estimators of Multiplicative Models for GEI. In Genotype by Environment Interaction, Manjit S. Kang, & HG. Gauch (eds), CRC Press. (pp. 200)
- Cotes, Jose M., J. Crossa, A. Sanches, and P.L. Cornelius. 2006. A Bayesian Approach for Assessing the Stability of Genotypes. <http://crop.scijournals.org/cgi/content/full/46/6/2654> [2 Juni 2008]
- Crossa J. 1990. Statistical Analysis of Multilocation Trials. Advances in Agronomy, 44 : 55-85.
- Croux C. & P. Filzmoser. 1998. Robust Factorization of a Data Matrix. Working paper. Online resources : <http://www.statistik.tuwien.ac.at/public/filz/papers/cstat98.ps>.
- Croux C., P. Filzmoser, G. Pison and P.J. Rousseeuw, 2003. Fitting Multiplicative Models by Robust Alternating Regressions. Statistics & Computing. Vol.13 No.1 p.23-36.
- de Falguerolles A. and Francis B. 1992. Algorithmic approaches for fitting bilinear models. COMPSTAT 1992, Proceedings in Computational Statistics, Vol. 1 (eds. Y. Dodge and J. Whittaker), Heidelberg, Physica-Verlag, pp. 77-82.

- de Falguerolles, A, 1996. Generalized Linear-Bilinear Models. An Abstract. Society of Computational Economics. 2nd International Conference on Computing and Finance. Genewa, Switzerland, 26–28 June 1996. <http://www.unige.ch/ce/ce96/defalgue/>
- Dhillon, I.S., and S. Sra. 2003. Modeling Data using Directional Distributions. Technical report. UTCS technical report.
- Edwards, Jode W. and Jean-Luc Jannink. 2006. Bayesian Modeling of Heterogeneous Error and Genotype Environment Interaction Variances. <http://crop.scijournals.org/cgi/content/full/46/2/820>. [11 Februari 2009]
- Efron B, Tibsirani R. 1993. An Introduction to the Bootstrap. New York: Chapman & Hall.
- Fisher, N.I., T.Lewis, and B.J.J. Embleton. 1987. Statistical Analysis of Spherical Data. Cambridge University Press. Cambridge.
- Gabriel K.R. 1978. Least squares approximation of matrices by additive and multiplicative models. Journal of the Royal Statistical Society B Vol. 40 No.2 p.186-196.
- Gabriel, K. R., 1998, Generalised Bilinear Regression. Biometrika. 85 (3):689-700.
- Gauch H.G. 1988. Model selection and validation for yield trial with interaction, Biometrics, 44:705-716.
- Gauch, H.G., Jr (1992) Statistical Analysis of Reginal Yield Traials: AMMI Analysis of Factorial Designs. Elsevier, Amsterdam, The Netherlands.
- Gilks, W.R., S. Richardson and D.J. Spiegelhalter. 1996. Marcov Chain Monte Carlo in Practice. Chapman & Hall. London.
- Gollob H.F. 1968. A Statistical Model which Combines Features of Factor Analytic and Analysis of Variance Techniques, Psychometrika, 33:73-116.
- Gower J. and Hand D. 1996. Biplots, Chapman & Hall, New York.
- Greenacre, M. J. 1984. Theory and Aplications of Correspondence Analysis. Academic Press. London.
- Groenen, P.J.F and Alex J. Koning. 2004a. A New Model for Visualizing Interactions in Analysis of Variance. Econometric Institute Report EI 2004-06.
- Groenen, P.J.F and Alex J. Koning. 2004b. Generalized Bi-additive Modelling for Categorical Data. Econometric Institute Report EI 2004-05.

- Hadi A. F., A. A. Mattjik, I. M. Sumertajaya, & H. Sa'diyah. 2008a. AMMI Model on Count Data: Model Log-Bilinear. The 3rd International Conference on Mathematics and Statistics (ICoMS-3). Moslem Statistician and Mathematician in South East Asia (MSMSEA). Bogor, Indonesia, 5-6 August 2008
- Hadi, A. F. & A. A. Mattjik, 2009a. Generalized AMMI Models For Assessing The Endurance of Soybean To Leaf Pest. The First International Seminar on Science and Technology (ISSTEC 2009). Universitas Islam Indonesia, Yogyakarta, 24 – 25 January 2009.
- Hadi, A.F, & A.A. Mattjik 2009b. Developing Robustness Of The AMMI Models By Robust Alternating Regression. Proceeding at The 4rd International Conference on Mathematics and Statistics. Bandar Lampung, August 2009.
- Hadi, A. F. & H. Sa'diyah, 2004. AMMI Model untuk Analisis Interaksi Genotip \times Lokasi. Jurnal Ilmu Dasar 1:33-41
- Hadi, A. F., H. Sa'diyah, & I. M. Sumertajaya. 2007. Penanganan Ketaknormalan Data pada Model AMMI menggunakan Transformasi BoxCox. Jurnal Ilmu Dasar, Vol 8. No 2. pp. 165-174.
- Hadi A. F, I. M. Sumertajaya, & I.M. Tirta, 2008b. Model AMMI pada Data Binomial: Model Logit-Bilinear. Seminar Nasional Sain dan Teknologi. Universitas Lampung, Bandar Lampung 17-18 Nopember 2008.
- Hoff, P. D. 2006. Model averaging and dimension selection for the singular value decomposition. <http://www.stat.washington.edu/research/reports/2006/tr494.pdf>. [16 Maret 2011].
- Jammalamadaka SR, SenGupta A. 1963. Topics in Circular Statistics (Series on multivariate analysis ; vol. 5). Singapore: World Scientific Publishing Co. Pte. Ltd.
- Jaya IGMN. 2008. Analisis interaksi genotipe x lingkungan menggunakan model persamaan struktural [Tesis]. Bogor: Sekolah Pasca Sarjana, Institut Pertanian Bogor.
- Johnson, RA & D.W. Wichern 1998. Applied Multivariate Statistical Analysis. 4th ed. Prentice Hall, New Jersey.
- Jolliffe, I T. 1986. Principal Component Analysis. Springer-Verlag. New York
- Kang MS. 2002. Genotype-Environment Interaction: Progress and Prospects. Di dalam : Kang MS, Editor. Quantitative Genetics, Genomic and Plant Breeding. Florida: CRC Pr. Hlm .221-243.

- Lavoranti OJ, et.al. 2007. Phenotypic stability via ammi model with bootstrap re-sampling. *Boletim de Pesquisa Florestal*. 2:45-52. http://www.lce.esalq.usp.br/tadeu/Osmir_Tadeu_Wojtek.pdf, [5 februari 2010].
- Liu, Genzhou. 2001. *Bayesian Computation for Linear-Bilinear Model*. University of Kentucky. Disertasi.
- Lynch, S.M. 2007. *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. Springer. NY.
- Mardia, K.V. & P.E. Jupp. 2000. *Directional Statistics*. Jonh Wiley & Sons Ltd. England.
- Mardia, K. V. & S. A. M. El-Atoum. 1976. Bayesian Inference for the Von Mises-Fisher Distribution. *Biometrika*, Vol. 63, no. 1. Hal. 203-206.
- Mathews P. 2000. The circular normal distribution. Ohio: Mathew Malnar and Bailey, Inc. <http://www.mmbstatistical.com/ToT/distofr.pdf>. [18 Mei 2010].
- Mattjik A. A. & Sumertajaya I. M. 2002. *Perancangan Percobaan dengan Aplikasi SAS dan MINITAB*. 2nd Ed. IPB Press. Bogor.
- Mattjik A. A., 2005. *Interaksi Genotipe dan Lingkungan dalam Penyediaan Sumberdaya Unggul*. Naskah Orasi Ilmiah Guru Besar Biometrika. Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Pertanian Bogor. Bogor.
- McCullagh, P. & Nelder, J. A. 1989. *Generalized Linear Models*. 2nd ed. Chapman and Hall, London.
- Mengersen, K.L., C.P. Robert, C. G. Jouyau. 1998. MCMC Convergence Diagnostics: A Review. <http://www.cvmcmc.eu/CntrlVrtsPapers/mengersen-robert.pdf> [08 Maret 2011].
- Morisson DF. 1978. *Multivariate Statistical Methods*, 2nd ed. Kogakusha: McGraw-Hill.
- Ntzoufras, I. 2009. *Bayesian Modeling Using WinBUGS*. John Wiley & Sons, Inc. New Jersey.
- Rencher AC. 2002. *Methods of multivariate analysis* 2nd ed. New York: John Wiley & Sons, Inc.
- Resende, M. D. P & R. Thompson, 2004. Factor Analytic Multiplicative Mixed Model in The Analysis of Multiple Experiments. *Rev. Mat. Estat.*, vol. 22. No. 2. Sao Paulo.

- Robert, C.P and G. Casella. 2010. *Introducing Monte Carlo Methods with R*. Springer Verlag. NY.
- Smith. A., B. Cullis, D. Lockett, G. Hollamby and R. Thompson. 2002. Exploring genotype-environment data using random effects AMMI models with adjustments for spatial field trend: Part 1-2: Theory-Applications. In: Kang, M.S. (Ed). *Quantitative Genetics, Genomics and Plant Breeding*. CAB International, UK
- Sues, E.A and B.E. Tumbro. 2010. *Introduction to Probability Simulation and Gibbs Sampling with R*. Springer. NY.
- Sumertajaya, IM. 2005. *Kajian Pengaruh Inter Blok dan Interaksi Pada Uji Lokasi Ganda dan Respon Ganda [Disertasi]*. Bogor: Program Pascasarjana IPB.
- Sumertajaya, I M. 1998. *Perbandingan Model AMMI dan Regresi Linier untuk Menerangkan Pengaruh Interaksi Percobaan Lokasi Ganda*. Tesis. Program Studi Statistika Sekolah Pascasarjana IPB, Bogor
- Terbeck W. and Davies P. 1998. Interactions and Outliers in The Two-way Analysis of Variance, *The Annals of Statistics*, 26:1279-1305.
- Turner, H. & D. Firth, 2009. *Generalized Nonlinear Models in R: An overview of the gnm package*. University of Warwick, UK.
- Van Eeuwijk, F A, 1995. Multiplicative Interaction in Generalized Linear Models. *Biometrics*, 51, p. 1017–1032.
- Viele, Kert. And C. Srinivasan. 1999. Parsimonious estimation of Multiplicative Interaction in Analysis of Variance using Kullback-Leiber Information. <http://sclab.yonsei.ac.kr/publications/Papers/DJ/E01002220577%5B1%5D.pdf> [16 Mei 2009].
- Yee, T.W. & T.J. 2003. Hastie, *Reduced-rank Vector Generalized Linear Models*. *Statistical Modelling*; 3: 15–41.
- Yuliyanti R. 2009. *Identifikasi genotipe yang memberikan kontribusi terhadap interaksi genotipe x lingkungan pada model ammi [Tesis]*. Bogor: Sekolah Pasca Sarjana, Institut Pertanian Bogor.