

PENDEKATAN METODE BAYES UNTUK PENDUGAAN PENGARUH INTERAKSI PADA MODEL AMMI

(Bayesian Approach for Estimating Interaction Effect of AMMI Model)

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

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

¹ Mahasiswa S2 Program Studi Statistika SPs IPB

² Departemen Statistika, FMIPA IPB

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 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 penelitian 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 ($\sigma_e^2 = 1$). 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 Ciharang) 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. $MSE = E\left((\hat{\delta} - \delta)^2\right) = \text{var}(\hat{\delta}) + Bias^2(\hat{\delta})$

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 $(\sigma_\mu^2, \sigma_{\tau_i}^2, \sigma_{\gamma_j}^2, \sigma_{\delta_{ij}}^2)$, 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 $\mu^{(l)}$ dari $\pi(\mu | \tau_i^{(l-1)}, \gamma_j^{(l-1)}, \delta_{ij}^{(l-1)}, \sigma^{2(l)})$

$$\pi(\mu | \tau_i, \gamma_j, \delta_{ij}, \sigma^2) \propto \exp \left\{ -\frac{rab\sigma_\mu^2 + \sigma^2}{2\sigma^2\sigma_\mu^2} \left(\mu - \frac{rab\sigma_\mu^2 \bar{y}_{\dots} + \sigma^2 \mu_\mu}{rab\sigma_\mu^2 + \sigma^2} \right)^2 \right\}$$

- b) Bangkitkan $\sigma^{2(l)}$ dari $\pi(\sigma^2 | \mu^{(l-1)}, \tau_i^{(l-1)}, \gamma_j^{(l-1)}, \delta_{ij}^{(l-1)})$

$$\pi(\sigma^2 | \mu, \tau_i, \gamma_j, \delta_{ij}) \propto IG \left[\frac{abr}{2} + \alpha_\sigma, \beta_\sigma + \frac{1}{2} \sum_{ijk} (y_{ijk} - \eta_{ij})^2 \right]$$

- c) Bangkitkan $\tau_i^{(l)}$ dari $\pi(\tau_i | \mu^l, \gamma_j^{(l-1)}, \delta_{ij}^{(l-1)}, \sigma^{2(l)})$

$$\pi(\tau_i | \mu, \gamma_j, \delta_{ij}, \sigma^2) \propto \exp \left\{ -\frac{rb\sigma_{\tau_i}^2 + \sigma^2}{2\sigma^2\sigma_{\tau_i}^2} \left(\tau_i - \frac{rb\sigma_{\tau_i}^2 \hat{\tau}_i + \sigma^2 \mu_{\tau_i}}{rb\sigma_{\tau_i}^2 + \sigma^2} \right)^2 \right\}$$

d) Bangkitkan $\gamma_j^{(l)}$ dari $\pi(\gamma_j | \mu^{(l)}, \tau_i^{(l)}, \delta_{ij}^{(l-1)}, \sigma^{2(l)})$

$$\pi(\gamma_j | \mu, \tau_i, \delta_{ij}, \sigma^2) \propto \exp \left\{ -\frac{ra\sigma_{\gamma_j}^2 + \sigma^2}{2\sigma^2\sigma_{\gamma_j}^2} \left(\gamma_j - \frac{ra\sigma_{\gamma_j}^2 \hat{\gamma}_j + \sigma^2 \mu_{\gamma_j}}{ra\sigma_{\gamma_j}^2 + \sigma^2} \right)^2 \right\}$$

e) Bangkitkan $\delta_{ij}^{(l)}$ dari $\pi(\delta_{ij} | \mu^{(l)}, \tau_i^{(l)}, \gamma_j^{(l)}, \sigma^{2(l)})$

$$\pi(\delta_{ij} | \mu, \tau_i, \gamma_j, \sigma^2) \propto \exp \left\{ -\frac{r\sigma_{\delta_{ij}}^2 + \sigma^2}{2\sigma^2\sigma_{\delta_{ij}}^2} \left(\delta_{ij} - \frac{r\sigma_{\delta_{ij}}^2 \hat{\delta}_{ij} + \sigma^2 \mu_{\delta_{ij}}}{r\sigma_{\delta_{ij}}^2 + \sigma^2} \right)^2 \right\}$$

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

$$*: \text{Persentase Perbaikan dugaan} = \frac{\text{MSE}_{\hat{\delta}_{ij}} - \text{MSE}_{\tilde{\delta}_{ij}}}{\text{MSE}_{\hat{\delta}_{ij}}} \times 100\%$$

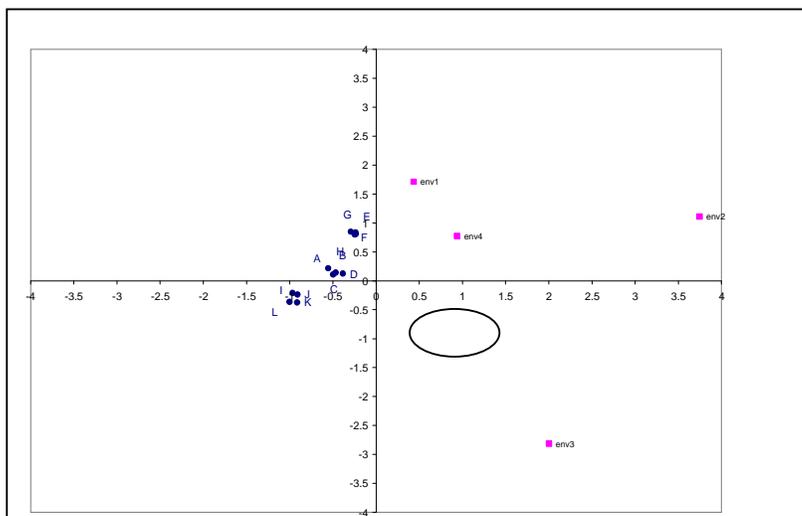
Pada Tabel 2, disajikan hasil simulasi klasifikasi genotipe menggunakan Biplot AMMI. Terlihat bahwa genotipe-genotipe yang diklasifikasikan stabil oleh metode Bayes, tidak terlalu berbeda dengan genotipe yang stabil dalam kondisi sesungguhnya (parameter). Sedangkan pada klasifikasi menggunakan MKT ada 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.

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

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

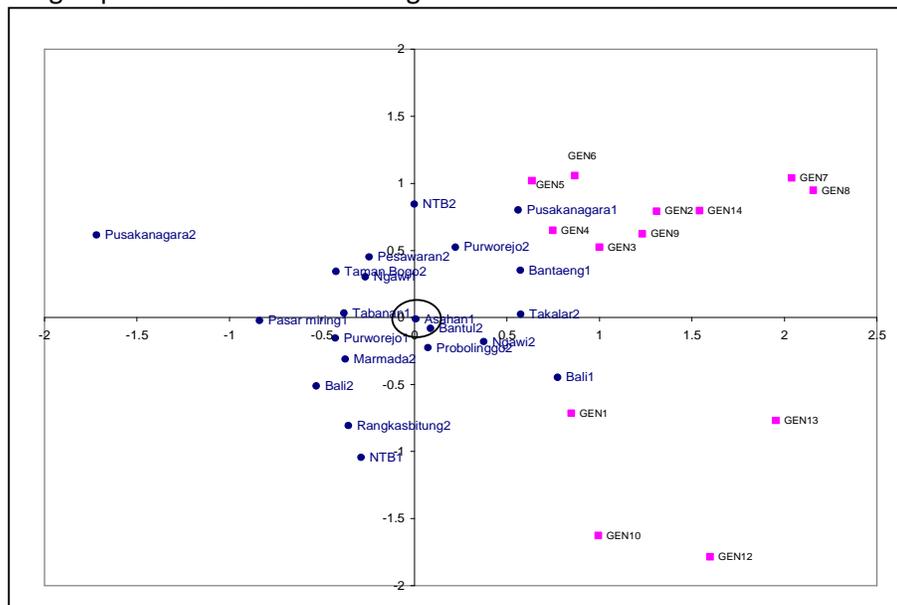
Perhitungan selang kepercayaan normal ganda pada taraf $\alpha = 0.05$ menghasilkan ellips dengan jari-jari panjang 0.47 dan jari jari pendek 0.38. Terlihat bahwa genotipe D

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

Data yang digunakan untuk ilustrasi berikut merupakan data percobaan tanaman padi BB Padi Sukamandi pada tahun 2008. Informasi prior untuk keragaman dan nilai tengah parameter model dihitung dari data tersebut.



Gambar 2. Biplot AMMI Data Padi

Pada Gambar 2 berikut disajikan Biplot AMMI dengan matriks pengaruh interaksi menggunakan pendugaan dengan pendekatan Bayes. Perhitungan selang kepercayaan normal ganda pada taraf $\alpha = 0.05$ menghasilkan ellips dengan jari-jari panjang 0.11 dan jari jari pendek 0.10. Terlihat pada bahwa tidak ada genotipe yang masuk ke dalam daerah kepercayaan ellips, yang berarti genotipe-genotipe tersebut dinyatakan sebagai genotipe yang tidak stabil.

Hasil biplot AMMI menggunakan penduga dengan metode MKT juga memberikan kesimpulan yang sama, dimana tidak ada genotipe yang dikategorikan stabil.

SIMPULAN

Hasil simulasi pendugaan pengaruh interaksi pada model AMMI menyatakan bahwa pendugaan dengan metode Bayes menghasilkan dugaan yang lebih baik, karena nilai MSE yang lebih kecil dibandingkan dengan dugaan pengaruh interaksi menggunakan metode MKT. Semakin besarnya keragaman lokasi, maka kemampuan metode Bayes memperbaiki dugaan cenderung meningkat. Sedangkan dengan semakin besarnya keragaman interaksi, kemampuan metode Bayes memperbaiki dugaan cenderung menurun.

Berdasarkan hasil simulasi Biplot AMMI untuk menentukan kestabilan genotipe, genotipe-genotipe yang dinyatakan stabil menggunakan pengaruh interaksi yang 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. 41/13.24.4/SPK/BG-PD/2009, 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 Statistician*. 46, 167-174. <http://www.istor.org/stable/2685208?origin=JSTOR-pdf> [29 Mei 2009].
- Berger, J.O.** 1985. *Statistical 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.