Forum Pascasarjana (1995) 18 (2) : 37 - 50.

ANALISIS PERAGAM GRUP DATA HETEROGEN[1])

(Covariance Analysis of Heterogenous Group Data)

Mohammad Masjkur, M. Sjarkani Musa, Aunuddin, dan Oetit Koswara[2])

ABSTRACT

This paper discussed the important of concomitant variables in regression analysis. Particularly, it was shown that without considering concomitant variables none of the combined regression models was reliable. In fact, the highest coefficient of determination R^2 was obtained from Mediterran soil group of data (65.8%). although R^2 for each specimen varied from 72.8% to 97.7%.

Considering K-soil contents as concomitant variables, it was shown that the coefficient of determination increased by 18 2% to 85.3%. The resulting R^2 for each model was ranging from 67.1% to 93.6%.

Another important finding from the research was the incorporation of concomitant variables resulted in more symmetric data distributions, smaller variances, and substantial increase in the regression slopes.

PENDAHULUAN

Gunawan (1984) telah melakukan suatu studi korelasi antar 10 metode ekstraksi kalium sebagai indeks ketersediaan hara kalium bagi tanaman jagung *(Zea mays* L.) dari enam jenis tanah. Sebanyak 46 spesimen tanah dari 46 tapak di pulau Jawa diidentifikasi masing- masing ada sebanyak 5, 11, 14, 5, 5, dan 6 spesimen termasuk jenis tanah Grumusol, Aluvial, Latosol, Mediteran, Podsolik, dan Regosol.

Spesimen-spesimen tanah dari suatu jenis tanah menurut klasifikasi tingkat jenis (great group) belum tentu memberikan trend respons yang sama terhadap pemupukan kalium. Sehingga dari suatu grup satuan-satuan percobaan heterogen penggunaan satu persamaan prediksi gabungan (yang dikehendaki dapat mewakili semua spesimen anggotanya) memerlukan evaluasi.

Timm (1975) mengemukakan prosedur statistika untuk menguji apakah dua atau lebih persamaan regresi berimpit, sejajar, atau tidak sama. Hal itu, berdasarkan anggapan bahwa kurva-kurva regresi mempunyai sabuk-sabuk kepercayaan yang sama. Anggapan tersebut tidak selalu terpenuhi oleh data empirik.

Faktor-faktor beragam yang tidak dikendalikan dalam rancangan percobaan, tetapi mungkin berperan dalam membangkitkan grup data heterogen, dapat digunakan sebagai pengendali (koreksi), yaitu sebagai peubah konkomitan atau kovariat.

---

[1]) Sebagian dari Tesis Magister Sains Jurusan Statistika Terapan, Fakultas Pascasarjana Institut Pertanian Bogor. Penelitian dibiayai oleh TMPD, Depdikbud.

[2]) Berturut-turut tiga staf pengajar Fakultas MIPA IPB, dan staf pengajar Fakultas Pertanian IPB

Pada tulisan ini akan disajikan teknik, apakah regresi-regresi dari S spesimen tanah dari suatu jenis tanah dapat diwakili hanya oleh satu regresi gabungan. Bilamana tidak dapat diperoleh suatu persamaan regresi gabungan yang mewakili, maka apakah penggunaan teknik kovarians efektif untuk suatu grup data heterogen.

ANALISIS DATA

Regresi **Polinomial** Ortogonal

Model regresi polinomial ortogonal ordo ke-k ($k \leq a-1$) adalah,

$$Y_{ij} = \beta_0 p_0(x_i)\ \beta_1 p_1(x_i) + .. + \beta_k p_k(x_i) + \varepsilon_{ij},\quad i=1,2,...,a;\quad j=1,2,...,r \qquad [1]$$

dengan anggapan $\varepsilon_{ij} \sim NIID(0,\sigma^2)$.
Sedangkan,

$$p'_g = [p_g(x_1), p_g(x_2), ..., p_g(x_a)],\qquad g = 1,2,...,k$$

Unsur-unsur p'_g telah didaftarkan sebagai koefisien-koefisien polinomial ortogonal untuk suatu faktor dengan taraf-taraf berjarak sama.

Regresi Untuk Suatu Jenis **Tanah**

Andaikan b_k ($k=0,1,2,...,a-1$) ialah penduga bagi βk. Dari suatu contoh acak berukuran n yang dianggap berasal dari suatu populasi obyek berukuran N, model,

$$y_i = b_0 p_0(x_i) + b_1 p_1(x_i) + ... + b_k p_k(x_i) + e_i,\qquad i=1,2,...,a \qquad [3]$$

misalnya, dinamakan model contoh. Untuk semua contoh (masing-masing berukuran n) berbeda yang mungkin, diperoleh populasi C(N,n) garis regresi contoh. Model,

$$y_i = \beta_0 p_0(x_i) + \beta_1 p_1(x_i) + ... + \beta_k p_k(x_i) + \varepsilon_i,\qquad i=1,2,...a \qquad [4]$$

dinamakan sebagai model populasi, yaitu dari populasi contoh- contoh acak berukuran n. Populasi heterogen terdiri atas beberapa anak populasi, yang masing-masing berukuran $N_1, N_2, ...,N_s$.

Jika terdapat s regresi dari suatu data heterogen yang terdiri atas s grup data, maka beberapa regresi mungkin berimpit (berprofil sama), sejajar (berprofil serupa), atau berprofil tidak sama. Penggabungan s regresi menjadi satu persamaan garis regresi yang dimaksudkan sebagai suatu persamaan dugaan bagi seluruh data s anak populasi belum tentu mewakili gugus data (Timm, 1975).

Analisis Peragam

Kegunaan analisis peragam adalah untuk mengendalikan galat yang terkandung dalam satuan-satuan percobaan, yang tidak terkendalikan atau terperhitungkan dalam rancangan percobaan (Cox dan McCullagh, 1982; Snedecor dan Cochran, 1956).

Andaikan x_{ijq} ialah nilai peubah konkomitan X yang terdapat pada satuan percobaan ke-j spesimen tanah ke-q, yang menerima perlakuan ke-i. Apabila rancangan percobaannya ialah suatu rancangan acak lengkap, maka model peragam bagi percobaan tersebut adalah,

$$y_{ijq} = \mu + \tau i + \beta (x_{ijq} - x_{...}) + \varepsilon_{ijq}, \; i=1,2,...,a; \quad j=1,2,...r, q=1,2,..., p \qquad [5]$$

Persamaan di atas dapat dicatat sebagai,

$$y_{ijq} - \beta(x_{ijq} - x_{...}) = \mu + \tau i + \varepsilon_{ijq} = pi + \varepsilon ijq \qquad [6]$$

sehingga,

$$y'_{ijq} = y_{ijq} - \beta (x_{ijq} - x_{...}) \qquad [7]$$

dapat disebut sebagai data respons dikoreksi, yaitu oleh nilai peubah konkomitan. Koefisien regresi β adalah suatu faktor pengganda dalam pengkoreksian.

HASIL DAN PEMBAHASAN

Regresi **Polinomial** Ortogonal

Tabel Lampiran 1 menyajikan koefisien-koefisien regresi (b_i) dan koefisien determinasi (R^2) model lengkap maupun reduksi, dan ragam galat regresi model lengkap serta reduksi untuk tiap spesimen tanah. Koefisien-koefisien regresi model reduksi dalam tabel tersebut, dapat dikenali dari bilangan yang dicetak tebal. Tabel juga menyajikan statistik-statistik yang sama untuk masing- masing jenis tanah, yaitu untuk seluruh spesirnen tanah dalam satu grup apriori berdasarkan jenis tanahnya.

Prosedur penyederhanaan model dilakukan menurut cara yang dianjurkan Allen dan Cady (1982). Suatu model reduksi dapat dipertimbangkan bilamana: (i) bias penyederhanaan model masih dapat ditenggang dan (ii) masih mempunyai derajat keterandalan cukup tinggi. Patokan yang digunakan penulis dalam hal perlu tidaknya pereduksian suatu model adalah koefisien determinasi (R^2) sebesar 70%.

Dari 46 percobaan, enam model lengkap diantaranya memiliki derajat keterandalan kurang memadai ($R^2 < 70\%$), yaitu: (a) spesimen dari Brebes dan Probolinggo dari jenis tanah Aluvial, (b) spesimen dari Sukabumi dari jenis tanah Podsolik, (c) spesimen dari

Majenang dari jenis tanah Grumusol. Hal ini mungkin disebabkan karena terdapat penyimpangan (ketaksesuaian) model. Pereduksian model dalam hal ini tidak relevan, karena model lengkapnya sendiri kurang terandalkan.

Dari 40 percobaan yang masing-masing mempunyai $R^2 > 70\%$, lima percobaan di antaranya tidak dapat direduksi karena pereduksian model menghasilkan R^2 kurang dari 70%. Sehingga model yang dipertimbangkan adalah model kuartik (ordo ke-4), yaitu untuk (a) spesimen dari Pasuruan dan Madiun dari jenis tanah Aluvial, (b) spesimen dari Purworejo dari jenis tanah Mediteran, dan (c) spesimen dari Gunung Muria dan Malang dari jenis tanah Regosol. Model-model yang masih dapat direduksi diberikan oleh:

(1) jenis tanah Aluvial sebanyak tujuh spesimen dengan ordo ke-1 sebanyak dua spesimen dan ordo ke-2 sebanyak lima spesimen,

(2) jenis tanah Podsolik sebanyak empat spesimen dengan ordo ke-1 sebanyak tiga spesimen dan ordo ke-2 satu spesimen,

(3) jenis tanah Latosol sebanyak tiga belas spesimen dengan ordo ke-1 sebanyak sepuluh spesimen dan ordo ke-2 sebanyak tiga spesimen,

(4) jenis tanah Mediteran sebanyak empat spesimen dengan ordo ke- 1 satu spesimen dan ordo ke-2 sebanyak tiga spesimen,

(5) jenis tanah Regosol sebanyak empat spesimen dengan ordo ke-1 sebanyak tiga spesimen dan ordo ke-2 satu spesimen,

(6) jenis tanah Grumusol sebanyak tiga spesimen dengan ordo ke-2 sebanyak dua spesimen dan ordo ke-3 satu spesimen.

Dengan respons berupa fungsi linear berarti bahwa penambahan tiap satuan pupuk K memberikan tambahan serapan K yang sama besar. Hal ini dimungkinkan sebab sesuai dengan ciri unsur K dalam hubungannya dengan tanaman, dimana K tersedia dalam tanah akan diserap oleh tanaman dalam jumlah yang melebihi kebutuhan optimumnya (Buckman dan Brady, 1964). Sedangkan bentuk respons dalam fungsi sampai komponen kuadratik, kubik, atau kuartik berarti bahwa penambahan tiap satuan pupuk K memberikan tambahan satuan serapan K yang tidak sama besar (pengaruh non linear).

Regresi Untuk Suatu Jenis Tanah

Hasil pengujian kehomogenan ragam-ragam dalam perlakuan dengan uji Bartlett untuk masing-masing percobaan tunggal dapat dilihat dalam Tabel Lampiran 2.

Dari 46 percobaan, terdapat lima percobaan yang ragam-ragamnya heterogen, yaitu spesimen dari Jombang (Aluvial), spesimen dari Bekasi (Podsolik), spesimen dari Majenang (Latosol), dan spesimen dari Pelabuhan Ratu dan Kebumen (Mediteran). Sedangkan untuk percobaan-percobaan lainnya ragam-ragamnya homogen.

Dengan demikian pada jenis tanah Aluvial, Podsolik, Latosol, dan Mediteran tidak dapat dilakukan pengujian keberimpitan atu kesejajaran garis-garis regresi secara formal menurut cara yang diajukan oleh Timm (1975). Walaupun untuk jenis tanah

Regosol, masing-tilasing regresi dari spesimen-spesimen memenuhi anggapan kehomogenan ragam-ragam, tetapi sabuk-sabuk kepercayaannya tidak sama, karena ragam gabungan untuk spesitnen-spesimen tidak sama. Sehingga juga tidak dapat dilakukan pengujian formal. Hanya pada jenis tanah Grumusol saja yang mungkin dilakukan pengujian keberinipitan atau kesejajaran regresi, karena anggapan bahwa regresi-regresi mempunyai sabuk-sabuk kepercayaan yang sama terpenuhi. Oleh karena itu, untuk selanjutnya keberimpitan atau kesejajaran regresi-regresi diperiksa secara visual saja.

Gambar I menyajikan garis-garis regresi spesimen-spesimen dalam suatu jenis tanah. Secara visual, untuk semua jenis tanah tampak adanya interaksi antara dosis pemupukan K dengan spesimen tanah. Hal tersebut ditunjukkan terdapatnya dua atau lebih garis regresi yang tak sejajar. Adanya interaksi antara dosis pemupukan K dengan spesimen tanah berarti bahwa penambahan atau perubahan sebesar satuan dosis pemupukan K terhadap spesitnen-spesimen tanah (dalam suatu jenis tanah) memberikan tambahan-tambahan respons (berupa K diserap tanaman) yang tidak sama besar.

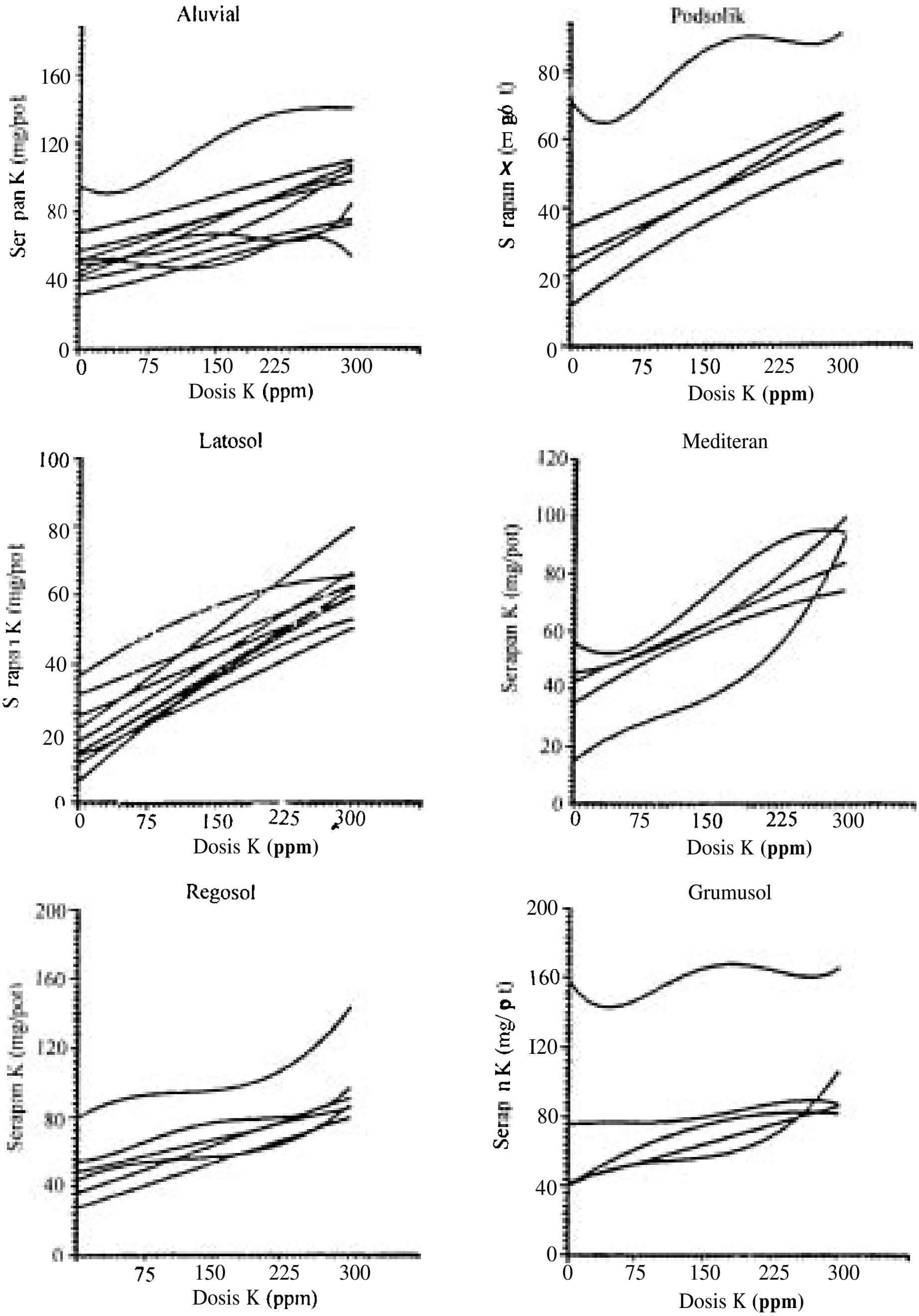
Gambar 2 menyajikan bagan kotak-garis dari respons spesimen- spesimen terhadap suatu dosis pemupukan K untuk masing-masing jenis tanah. Batas bawah garis menunjukkan kuantil $P_{2.5}$, sedangkan batas atas garis adalah kuantil $P_{97.5}$. Jadi, bagan kotak-garis yang disajikan ialah suatu sabuk 95%. Titik-titik di luar salah satu atau kedua batas tadi dapat dianggap sebagai pencilan-pencilan. Apabila lokasi median-median respons dari masing-masing jenis tanah dihubungkan, maka yang diperoleh ialah suatu garis regresi median dari regresi gabungan spesimen- spesimen dalam suatu grup (jenis) tanah.

Sebaran data respons percobaan gabungan pada tiap taraf dosis pemupukan kalium pada semua jenis tanah umumnya menjulur kekanan. Pada jenis tanah Latosol terdapat data pencilan. Wilayah antar kuartil (Q3-Q1) sebagai ukuran keragaman dalam perlakuan beragam antar dosis pemupukan kalium. Garis regresi median menunjukkan kecenderungan menaik, yang berarti bahwa serapan kalium meningkat dengan meningkatnya dosis pemupukan kalium.

Analisis Peragam

Informasi yang diberikan dalam Tabel Lampiran 1, dan Gambar I serta 2 ialah tanpa mempertimbangkan peubah konkomitan. Padahal kandungan K-tanah dari spesimen-spesimen berdasarkan hasil ekstraksi K-tanah menurut metode 1 M $CaCl_2$ cukup beragam. Ekstraksi dilakukan sebelum pemberian dosis-dosis K dalam percobaan pot.

Dapat diamati bahwa tidak satupun regresi-regresi yang mempunyai derajat keterandalan yang memadai. Koefisien determinasi terbesar diberikan oleh jenis tanah Mediteran (65.8%). Kurang terandalnya model regresi disebabkan karena koefisien keragamannya relatif besar, mengingat bahwa spesitnen-spesimen yang tergolong dalam suatu jenis tanah disamping memiliki kesamaan juga memiliki perbedaan-perbedaan dalam ciri fisik dan kimiawi, khususnya beragamnya kandungan K-tanah.



Gambar 1. Respons Spesimen-spesimen Dalam Jenis Tanah Terhadap Dosis Pemupukan K

Figure 1. Response of the soil speciments to Potassium.

Sehingga, regresi-regresi mungkin dapat dikoreksi dengan satu atau beberapa peubah konkomitan yang relevan dan tersedia, dalam data percobaan ini ialah kandungan K-tanah tersedia.

Hasil analisis regresi dengan peragam untuk masing-masing jenis tanah diringkaskan seperti dapat dilihat dalam Tabel 1 berikut:

Tabel 1. Ringkasan Hasil Analisis Peragam Grup Data
Table 1. Summary of anova of data group

Grup Data	Koefisien-koefisien regresi (Polinomial ortogonal)					R^2 Model		Proporsi keragaman (model lengkap)			Koefisien regresi peubah konkomitan
	b0	Lin	Kuad	Kub	Kuar	Lengkap	Reduksi	Pupuk K	K-tanah	Galat	
Aluvial	21,5	9,7	0,1	-0,4	0,1	67,1	67,0	26,6	40,5	32,9	0,2506
Podsolik	18,8	9,1	-0,4	-0,2	0,2	88.2	88,0	36,1	52,1	11,8	0,2049
Latosol	21,7	10,3	0,2	0,3	-0,2	74,6	74,5	31,9	42.7	25.4	0,2235
Mediteran	34.6	12.6	1,6	0,3	-0,1	84,0	83.9	65,8	18.2	16.0	0,1731
Regosol	24.0	11,5	1,5	1,2	-0,1	87,3	86,8	45.5	41,8	12.8	0,2106
Grumusol	37,5	8,6	0,5	0,5	0,2	93,6	93,5	9,4	84.2	6,4	0,1528

Catatan: Yang dicetak tebal adalah koefisien-koefisien regresi model reduksi

Dari tabel tersebut dapat diketahui bahwa:

(1) Koefisien determinasi (R^2) regresi dengan peubah konkomitan lebih besar dari R^2 tanpa peubah konkomitan, dengan perbedaan berkisar dari 18,2% (Mediteran) sampai dengan 85,3% (Grumusol),

(2) Pereduksian model dapat dilakukan karena model reduksi masih mempunyai derajat keterandalan cukup tinggi dan tidak berbeda jauh dari model lengkapnya,

Pereduksian model untuk masing-masing jenis tanah menunjukkan bahwa jenis tanah Aluvial, Podsolik, Latosol, dan Grumusol berespons linear (ordo ke-1). Sedangkan jenis tanah Mediteran dan Regosol berespons kuadratik (ordo ke-2).

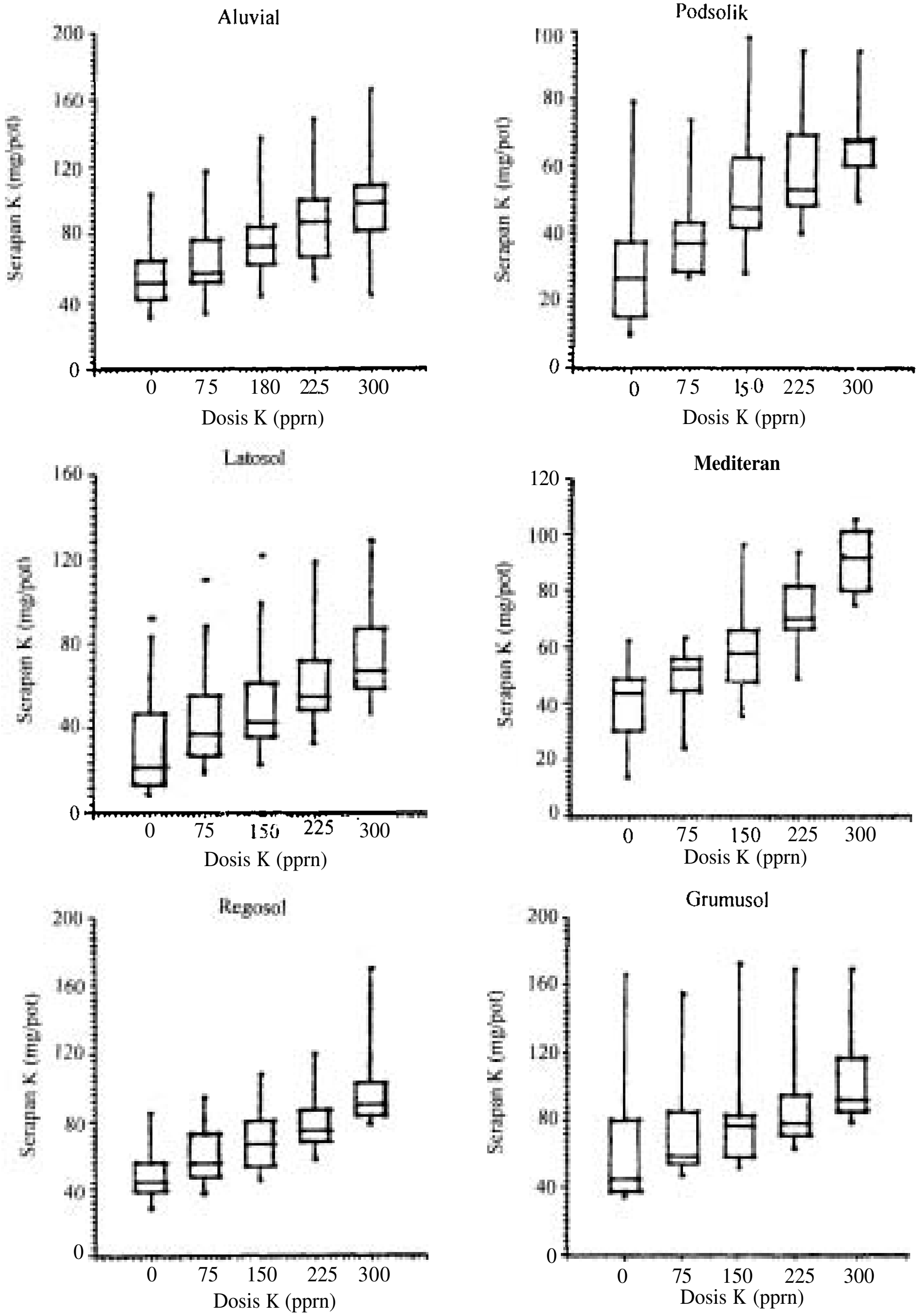
Gambar 3 menyajikan trend respons jenis tanah terhadap dosis pemupukan kalium dengan mempertimbangkan K-tanah sebagai peubah konkomitan. Bilamana informasi yang disajikan dalam Gambar 3 (dengan peubah konkomitan) dibandingkan dengan yang diperoleh dari Gambar 2 (tanpa peubah konkomitan), maka agaknya dapat disimpulkan bahwa:

(1) sebaran data respons percobaan gabungan dengan K-tanah sebagai peubah konkomitan makin simetrik,

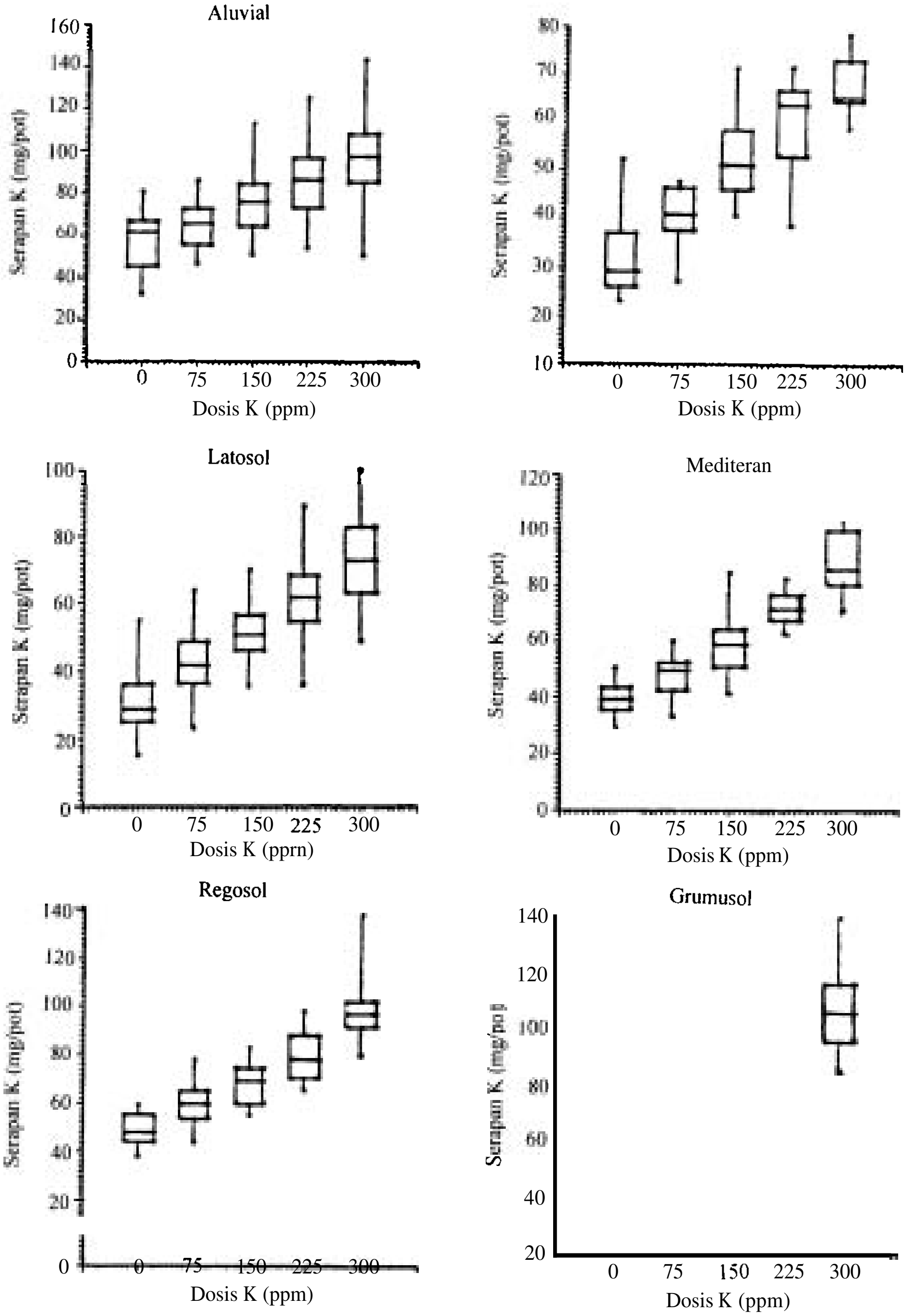
(2) tidak ada lagi data pencilan,

(3) wilayah antar kuartil (Q3-Q1) makin sempit, walaupun relatif masih beragam pada taraf-taraf dosis pemupukan kalium,

(4) garis regresi median naik cukup tajam dengan meningkatnya dosis pemupukan kalium.



Gambar 2. Plot Kuantil untuk Sabuk 95% dan Garis Median untuk Regresi Grup Data
Figure 2. Plot of Quantil at 95% confidence interval and median lines for regression of data group



Gambar **3.** Plot Kuantil untuk Sabuk 95% dan Garis Median untuk Regresi Grup Data dengan K-tanah sebagai Peubah Konkomitan

Figure **3.** Plot of quantie at 95% confidence interval and median lines for regression of data group with soil-K as Konkomitan variable

KESIMPULAN

(1) Tanpa mempertimbangkan peubah konkomitan, tidak satupun regresi gabungan yang dapat diandalkan. Koefisien determinasi terbesar diberikan grup data jenis tanah Mediteran (65,8%), walaupun R^2 dari masing-masing spesimen berkisar dari 72,8% sampai dengan 97,7%.

(2) Dengan mempertimbangkan K-tanah sebagai peubah konkomitan dalam analisis regresi diperoleh bahwa:

(a) koefisien determinasi R^2 meningkat berkisar dari 18,2% sampai dengan 85,3%. R^2 yang diperoleh berkisar dari 67,1% sampai dengan 93,6%,

(b) dengan K-tanah sebagai peubah konkomitan, maka sebaran data makin simetrik, tidak ada lagi data pencilan, wilayah antar kuartil (Q3-Q1) makin sempit, dan garis regresi median menaik makin tajam.

DAFTAR PUSTAKA

Allen, D. M. and F. B. Cady. 1982. Analyzing Experimental Data by Regression. Lifetime Learning Publications. A Division of Wadsworth, Inc., California.

Buckman, H. O. and N. C. Brady. 1964. The Nature and Properties of Soils. Sixth Edition. McMillan Co., New York.

Cox, D. R. and P. McCullagh. 1982. Some aspects of analysis of covariance. Biometrics 38: 541-561.

Gunawan. 1984. Studi korelasi berbagai rnetode ekstraksi kaliurn sebagai indeks ketersediaan hara kalium bagi tanaman jagung (*Zea mays* L.) pada berbagai jenis tanah di pulau Jawa. Jurusan Ilmu- ilrnu Tanah, Fakultas Pertanian. Institut Pertanian Bogor. Bogor. Tesis tidak dipublikasikan.

Snedecor, G. W. and W. G. Cohran. 1956. Statistical Methods. Fifth Edition. The Iowa State University, Iowa.

Timm, N. H 1975 Multivariate Analysis; With applications in education and psychology. Wadsworth Publishing Co. Inc., Belmont, California, hal : 331-347.