

## ANALISIS MULTIVARIAT PADA DATA INDEKS GEOMAGNET GLOBAL

<sup>1</sup>John Maspupu

<sup>1</sup>Pussainsa LAPAN, Jl. Dr. Djundjuna No. 133 Bandung 40173,  
Tlp. 0226012602 Pes. 106. Fax. 0226014998  
E-mail: [john\\_mspp@yahoo.com](mailto:john_mspp@yahoo.com)

**Abstrak.** Makalah ini membahas suatu analisis multivariat yang diterapkan pada data indeks  $\sum Kp$  dari beberapa tempat observasi (stasion geomagnet-SG). Indeks  $\sum Kp$  yang dimaksud dalam pembahasan ini adalah penjumlahan indeks  $Kp$  (indeks geomagnet global) untuk setiap pengamatan dalam satu hari. Perincian dari indeks  $\sum Kp$  yang merupakan bilangan cacah ini dapat dijelaskan sebagai berikut, untuk  $0 \leq \sum Kp \leq 20$  ditafsirkan sebagai tingkat gangguan geomagnet yang sangat rendah. Kemudian untuk  $21 \leq \sum Kp \leq 52$  ini mengindikasikan tingkat gangguan geomagnet yang sedang. Selanjutnya untuk  $53 \leq \sum Kp \leq 72$ , ini berarti tingkat gangguan geomagnet yang sangat tinggi. Konsep analisis multivariat ini juga merupakan suatu teknik atau metode dependensi yang biasanya digunakan untuk meramalkan kondisi-kondisi data variabel tak bebas bersifat non metrik atau kualitatif, berdasarkan data prediktornya yaitu variabel bebas yang bersifat metrik atau kuantitatif. Dengan demikian tujuan pembahasan makalah ini adalah untuk meramalkan kondisi kualitatif geomagnet (terganggu atau tidak terganggu), dengan menggunakan data kuantitatif indeks  $\sum Kp$  dari beberapa stasion observasi geomagnet serta dua lokasi observasi yang memiliki data harian bilangan bintang matahari. Kedua data terakhir ini merupakan data pendukung indeks geomagnet global. Hasil aplikasi analisis diskriminan pada data-data di atas menunjukkan bahwa kondisi kualitatif geomagnet di lokasi atau sekitar lokasi-lokasi tersebut tidak terganggu. Selain itu kontribusi dari hasil aplikasi ini, juga dapat memberikan informasi tentang kondisi geomagnet pada setiap pengamatan secara global.

**Kata kunci :** Analisis multivariat, Data indeks  $\sum Kp$ .

### 1. Pendahuluan

Analisis multivariat (*multivariate analysis*) merupakan suatu kelompok dari teknik atau metode ketergantungan (*dependence methods*). Konsep analisis multivariat ini awalnya digunakan oleh Miller pada tahun 1962 dalam kasus-kasus peramalan cuaca. Selain itu beberapa aplikasi dari konsep analisis multivariat untuk kasus-kasus peramalan badai topan juga telah dipublikasikan oleh Lehmler dan kawan-kawannya pada tahun 1997. Begitu pula peran konsep analisis multivariat untuk kasus-kasus peramalan lainnya dapat dibaca pada makalah Lawson, M.P. & Cervený, R.S., (1985) dan Ward, M.N. and Folland, C.K., (1991). Oleh karena itu dengan mempertimbangkan beberapa referensi yang telah dikemukakan di atas, muncul pemikiran untuk mengaplikasikan analisis multivariat ini pada data indeks geomagnet global. Dengan demikian tujuan pembahasan makalah ini adalah untuk meramalkan kondisi kualitatif geomagnet

(*terganggu* atau *tidak terganggu*), dengan menggunakan data kuantitatif indeks  $\sum Kp$  dari beberapa stasion observasi geomagnet. Namun yang menjadi masalah adalah bagaimana proses peramalan kondisi kualitatif geomagnet ini dilakukan? Atau bagaimana prosedur penerapan analisis multivariat untuk kasus peramalan kondisi geomagnet tersebut? Kemudian informasi apa saja yang diperoleh dari hasil analisis multivariat kasus ini? Untuk menjawab semua permasalahan di atas ini, perlu disusun suatu metodologi yang tepat serta dapat memberikan solusi secara tuntas dan bermanfaat.

## 2. Metodologi

Konsep yang digunakan dalam pembahasan makalah ini adalah menyangkut analisis multivariat dengan satu variabel tak bebas yaitu kondisi geomagnet yang bersifat kualitatif (*nonmetrik*) dan terdiri dari dua kategori (*terganggu* atau *tidak terganggu*). Sedangkan lima variabel bebasnya bersifat kuantitatif (*metrik*) yaitu tiga lokasi observasi yang memiliki data indeks  $\sum Kp$  serta dua lokasi observasi yang memiliki data harian bilangan bintang matahari. Selanjutnya tahapan analisis multivariat ini dapat dijabarkan dalam beberapa langkah berikut :

- i). Kompilasi data pengamatan (data asli atau data simulasi) dan merumuskan masalah multivariat yaitu dengan mengidentifikasi : tujuan, variabel bebas (*predictor*) dan variabel tak bebas (*dependent*). Kemudian pembagian sampel data menjadi duabagian yaitu:
  - a. sampel estimasi (*estimation sample*) yang digunakan sebagai dasar pembuatan estimasi koefisien fungsi multivariat.
  - b. sampel validasi (*validation sample*) yang digunakan untuk validasi. Penjelasan terperinci tentang dua jenis sampel ini dapat dilihat pada makalah Maspupu, J., (2011).

- ii). Membuat estimasi meliputi pengembangan suatu kombinasi linier

$$M_i = b_0 + \sum_{j=1}^k b_j x_{ij}$$

dari prediktor yang biasanya disebut sebagai fungsi multivariat

$M_i$ , sehingga kelompok sedapat mungkin sangat berbeda berdasarkan nilai-nilai prediktor. Penjelasan terperinci dapat dilihat pada makalah Maspupu, J., (2011).

- iii). Penentuan signifikansi fungsi multivariat melibatkan nilai signifikan  $S$  yang dihitung dari tabel ANOVA satu arah dengan pengelompokkan variabel sebagai variabel bebas yang kualitatif.  $S = P ( S \geq S_0 ) = p$  dengan  $H_0 : B_1 = B_2 = \dots = B_k = 0$  dan  $H_1 : B_j \neq 0$ . Keputusannya : Jika  $p \leq \alpha$  maka  $H_0$  ditolak dan jika  $p > \alpha$  maka  $H_0$  diterima. Atau dapat juga menggunakan tabel F dengan keputusannya sebagai berikut:

Jika  $F_H \geq F_{\alpha, (v_1, v_2)}$  maka  $H_0$  ditolak dan jika  $F_H < F_{\alpha, (v_1, v_2)}$  maka  $H_0$  diterima. Dalam hal ini  $\alpha$  adalah tingkat keberartian (*level of significance*) dan  $(v_1, v_2)$  adalah derajat kebebasan (*degree of freedom*) distribusi F. Beberapa ungkapan yang terkait analisa multivariat ini dinyatakan dalam beberapa definisi yang dapat dilihat pada makalah Maspupu, J., (2011). Dalam analisis multivariat perlu diasumsikan bahwa, setiap kelompok merupakan suatu sampel dari populasi normal

- multivariat (*multivariate normal population*) dan setiap populasi mempunyai matriks kovarian yang sama.
- iv). Interpretasi hasil perhitungan koefisien fungsi multivariat ini serupa seperti menganalisis regresi linier berganda. Jika terjadi kolinieritas berganda (*multicollinearity*) didalam prediktor maka sulit untuk menarik kesimpulan bahwa apakah ada perbedaan rata-rata antar kelompok. Untuk menghindari terjadinya kolinieritas ganda beberapa pendekatan dapat dilakukan dengan memperhatikan atau melihat makalah Dillon, W.R. & Goldstein, M., (1984) dan Tabachnick, B.G. & Fidell, L.S., (2007).
- v). Penilaian terhadap validitas analisis multivariat mencakup pengembangan matriks klasifikasi. Penjelasan akurasi klasifikasi dikatakan tercapai dengan analisis multivariat multivariat jika lebih besar 25 % dari pada dengan cara kebetulan (lihat Dillon, W.R. & Goldstein, M., (1984) atau Tabachnick, B.G. & Fidell, L.S., (2007)). Penjelasan tentang ungkapan rasio kesuksesan dan bobot multivariat serta cara pendekatan untuk memperkirakan koefisien fungsi multivariat dapat dilihat pada makalah Maspupu, J., (2011).

### 3. Hasil dan Pembahasan

Data yang digunakan dalam pembahasan makalah ini adalah data-data indeks  $\sum Kp$  dari tiga lokasi SG (Stasiun Geomagnet). Data-data ini diamati selama 8 selang waktu dalam sehari dengan pengertian tiap selang waktu adalah 3 jam. Juga menggunakan data pendukung lainnya yaitu data harian bilangan bintang matahari  $R_h$  dari dua lokasi SPM (Stasiun Pengamat Matahari). Sampel data-data tersebut diambil dari 42 hasil observasi yang didistribusi dalam dua jenis sampel masing-masing terdiri dari 30 obyek atau kasus untuk sampel estimasi dan 12 obyek atau kasus untuk sampel validasi. Kemudian dengan menerapkan langkah i) dari bagian metodologi tersebut akan diperoleh hasil kompilasi data indeks  $\sum Kp$  dari tiga lokasi SG dan  $R_h$  dari dua lokasi SPM seperti yang ditabulasikan dalam Tabel 1 dan Tabel 2.

**Tabel 1.** Data harian sebagai sampel untuk estimasi dari  $R_h$  di dua lokasi observasi dan indeks  $\sum Kp$  di tiga lokasi observasi.

Observasi ke- n	$R_h$ di Lokasi SPM 1 $X_1$	indeks $\sum Kp$ di Lokasi SG <sub>1</sub> $X_2$	indeks $\sum Kp$ di Lokasi SG <sub>2</sub> $X_3$	indeks $\sum Kp$ di Lokasi SG <sub>3</sub> $X_4$	$R_h$ di Lokasi SPM 2 $X_5$
1.	55	3	8	8	52,7
2.	68	2	5	7	75,0
3.	62	2	2	5	46,2
4.	51	3	7	5	57,0
5.	43	2	7	6	50,2
6.	61	3	5	7	70,3
7.	52	4	6	3	62,9
8.	36	6	7	4	48,5
9.	57	5	7	5	64,1

10.	45	4	6	6	68,1
11.	44	3	5	7	73,4
12.	64	2	4	8	71,9
13.	55	3	2	8	36,2
14.	57	5	5	2	43,2
15.	37	4	6	6	50,4
16.	42	5	6	4	44,1
17.	54	3	5	3	56,2
18.	56	6	1	5	49,3
19.	58	4	4	2	62,0
20.	58	5	5	6	32,1
21.	56	3	5	6	41,8
22.	36	2	8	2	57,0
23.	50	2	6	5	33,4
24.	48	4	3	4	37,5
25.	42	3	3	7	41,3
26.	45	5	6	1	38,3
27.	57	4	1	3	55,0
28.	51	6	3	8	46,1
29.	64	3	6	2	35,0
30.	54	2	2	3	37,3

**Tabel 2.** Data harian sebagai sampel untuk validasi dari  $R_h$  di dua lokasi observasi dan indeks  $\sum Kp$  di tiga lokasi observasi.

Observasi ke- n	$R_h$ di Lokasi SPM 1 $X_1$	indeks $\sum Kp$ di Lokasi SG <sub>1</sub> $X_2$	indeks $\sum Kp$ di Lokasi SG <sub>2</sub> $X_3$	indeks $\sum Kp$ di Lokasi SG <sub>3</sub> $X_4$	$R_h$ di Lokasi SPM 2 $X_5$
1.	46	3	4	7	68,1
2.	56	7	7	4	62,1
3.	54	4	6	7	35,0
4.	39	3	5	4	49,6
5.	44	6	6	6	39,4
6.	51	3	5	6	37,0
7.	37	4	4	3	54,5
8.	49	5	5	3	38,2
9.	55	3	6	5	50,8
10.	58	5	2	6	63,6
11.	60	3	7	3	54,0
12.	45	3	2	2	45,0

Dengan menerapkan langkah ii) dari bagian metodologi ini, diperoleh hasil analisis multi- variat untuk rata-rata kelompok dan simpangan baku dalam kelompok seperti yang tertera pada Tabel 3 dan Tabel 4.

**Tabel 3.** Hasil analisis multivariat untuk rata-rata kelompok.

Kelompok	$R_h$ di SPM 1	$\sum Kp$ di $SG_1$	$\sum Kp$ di $SG_2$	$\sum Kp$ di $SG_3$	$R_h$ di SPM 2
A (terganggu)	52,4667	3,4000	5,4667	5,8000	58,0067
B (tidak terganggu)	51,4000	3,8000	4,2667	4,0667	44,2667
Total	51,9333	3,6000	4,8667	4,9333	51,2167

**Tabel 4.** Hasil analisis multivariat untuk simpangan baku dalam kelompok.

Kelompok	$R_h$ di SPM 1	$\sum Kp$ di $SG_1$	$\sum Kp$ di $SG_2$	$\sum Kp$ di $SG_3$	$R_h$ di SPM 2
A (terganggu)	9,7165	1,2421	1,7674	1,8205	12,2630
B (tidak terganggu)	7,5668	1,3732	2,0517	2,0517	9,4844
Total	8,5740	1,3025	1,9780	2,0998	12,7952

Juga dengan memperhatikan definisi 3 dan definisi 4 pada makalah Maspupu, J., (2011), akibatnya diperoleh hasil perhitungan koefisien fungsi multivariat kanonik yang dibakukan dan tidak dibakukan seperti yang telah dicantumkan pada Tabel 5.

**Tabel 5.** Koefisien fungsi multivariat kanonik yang dibakukan dan tidak dibakukan.

Prediktor	Koefisien variabel bebas yang dibakukan	Koefisien variabel bebas yang tidak dibakukan
$R_h$ di SPM 1	0,21	$0,2454 \cdot 10^{-1}$
$\sum Kp$ di $SG_1$	0,47	0,4274
$\sum Kp$ di $SG_2$	0,10	$0,4964 \cdot 10^{-1}$
$\sum Kp$ di $SG_3$	0,23	0,1203
$R_h$ di SPM 2	0,74	$0,8477 \cdot 10^{-1}$
Konstanta	---	-7,9755

Begitupun dengan memperhatikan definisi 6 pada makalah Maspupu, J., (2011) , akibatnya diperoleh hasil perhitungan nilai Wilk's dan  $F_{1,28}$  untuk setiap variabel bebas seperti yang telah dicantumkan pada Tabel 6.

**Tabel 6.** Nilai Wilk's atau  $\mu$  statistik dan  $F_{1,28}$  untuk setiap variabel bebas (prediktor).

Prediktor	Nilai Wilk's	Nilai $F_{1,28}$	Signifikansi
$R_h$ di SPM 1	0,9544	1,34	0,2572
$\sum Kp$ di $SG_1$	0,6567	14,64	0,0007
$\sum Kp$ di $SG_2$	0,9248	2,28	0,1425
$\sum Kp$ di $SG_3$	0,8238	5,99	0,0209
$R_h$ di SPM 2	0,4531	33,80	0,0000

Menurut perhitungan dari definisi 1 pada makalah Maspupu, J., (2011) , akhirnya diperoleh hasil-hasil seperti yang tercantum dalam Tabel 7.

**Tabel 7.** Fungsi multivariat kanonik yang dievaluasi pada rata-rata kelompok.

Kode kelompok	Nilai fungsi
A atau 1	$\overline{D_A} = 1,29$
B atau 2	$\overline{D_B} = -1,29$

Selanjutnya dengan memperhatikan definisi 7 pada makalah Maspupu, J., (2011) , akibatnya diperoleh hasil perhitungan matriks korelasi dalam kelompok seperti yang telah dicantumkan pada Tabel 8.

**Tabel 8.** Matriks korelasi dalam kelompok.

Prediktor	$R_h$ di SPM 1	$\sum Kp$ di $SG_1$	$\sum Kp$ di $SG_2$	$\sum Kp$ di $SG_3$	$R_h$ di SPM 2
$R_h$ di SPM 1	1,0000				
$\sum Kp$ di $SG_1$	-0,4301	1,0000			
$\sum Kp$ di $SG_2$	-0,1971	-0,0168	1,0000		
$\sum Kp$ di $SG_3$	0,0174	0,0705	0,0843	1,0000	
$R_h$ di SPM 2	-0,0143	0,0889	0,1974	0,9148	1,0000

Selanjutnya untuk menentukan prosentase obyek atau kasus kelompok secara tepat, ini dapat diklasifikasi dalam Tabel 9 dan Tabel 10.

**Tabel 9.** Hasil klasifikasi kasus-kasus yang diseleksi dalam analisis estimasi.

Kode kelompok	Banyak kasus	Yang diprediksi	Anggota kelompok
A atau 1	15	12 80 %	3 20 %
B atau 2	15	0 0 %	15 100 %

**Tabel 10.** Hasil klasifikasi kasus-kasus yang tidak diseleksi dalam analisis validasi.

Kode kelompok	Banyak kasus	Yang diprediksi	Anggota kelompok
A atau 1	6	4 66,7 %	2 33,3 %
B atau 2	6	0 0 %	6 100 %

Dari Tabel 5 terlihat bahwa koefisien fungsi multivariat  $b_i$  untuk variabel bebas  $X_i$  ( $i = 1, \dots, 5$ ) yang telah dibakukan semuanya positif dengan perincian sebagai berikut :  $b_1 = 0,21$  ;  $b_2 = 0,47$  ;  $b_3 = 0,10$  ;  $b_4 = 0,23$  ; dan  $b_5 = 0,74$ . Ini berarti semua prediktor sebanyak 5 (lima) buah yang tercantum didalam fungsi multivariat  $M$  mempunyai pengaruh positif. Dengan demikian fungsi multivariat untuk variabel bebas yang telah

dibakukan adalah sebagai berikut:  $M = 0,21 X_1 + 0,47 X_2 + 0,10 X_3 + 0,23 X_4 + 0,74 X_5$ . Selain itu dari Tabel 5 juga terlihat bahwa koefisien fungsi multivariat  $b_i$  untuk variabel bebas  $X_i$  yang tidak dibakukan semuanya positif kecuali faktor konstanta  $b_0$  negatif. Dengan perincian sebagai berikut:  $b_1 = 0,02$ ;  $b_2 = 0,43$  ;  $b_3 = 0,05$  ;  $b_4 = 0,12$  dan  $b_5 = 0,08$  serta  $b_0 = -7,97$ . Sehingga dapat dikatakan bahwa fungsi multivariat untuk variabel bebas yang tidak dibakukan adalah:  $M = -7,97 + 0,02 X_1 + 0,43 X_2 + 0,05 X_3 + 0,12 X_4 + 0,08 X_5$ . Setelah dilakukan pengujian ternyata hanya ada 3 (tiga) dari 5 variabel bebas tersebut yang pengaruhnya sangat signifikan yaitu ,  $X_5$ ,  $X_4$  dan  $X_2$  . Ini dikarenakan nilai signifikansi dari masing-masing variabel tersebut adalah sebesar 0,000 ; 0,021 dan 0,001 yang semuanya kurang dari 5 % (lihat Tabel 6). Sehingga fungsi multivariat untuk variabel bebas yang telah dibakukan menjadi  $M = 0,47 X_2 + 0,23 X_4 + 0,74 X_5$  sudah cukup untuk melakukan suatu pengelompokan. Karena  $n_A = n_B = 15$  dan menurut Tabel 8,

$$\overline{D_A} = \overline{D_1} = 1,29 \quad \text{juga} \quad \overline{D_B} = \overline{D_2} = -1,29 \quad \text{maka nilai pemisah} \quad D_0 = \frac{\overline{D_A} + \overline{D_B}}{2} = \frac{1,29 - 1,29}{2} = 0 .$$

Selain itu hasil perhitungan fungsi multivariat untuk setiap data pengamatan dari beberapa lokasi seperti yang dicantumkan dalam Tabel 1 selalu memberikan nilai positif ( $> 0$  ). Sehingga menurut kriteria pada makalah Maspupu, J., (2011), obyek-obyek tersebut termasuk dalam kelompok B yaitu kondisi tidak terganggu. Selanjutnya prosentase obyek secara tepat untuk estimasi (lihat Tabel 9) adalah sebesar  $\frac{12+15}{30} = 0,90$ . Jadi prosentase kasus yang dikelompok secara tepat dalam analisa estimasi adalah 90 % . Sementara untuk validasi (lihat Tabel 10) adalah sebesar  $\frac{4+6}{12} = 0,833$ . Dengan demikian prosentase kasus yang dikelompok secara tepat dalam analisa validasi adalah 83,3 % . Menurut definisi *hit ratio* pada langkah v) di bagian metodologi dari makalah Maspupu, J., (2011), rasio suksesnya secara kebetulan adalah sebesar  $\frac{1}{2}$  atau 50 % . Jadi perbaikan yang dicapai dibandingkan dengan cara kebetulan adalah sebesar  $83,3 \% - 50 \% = 33,3 \% > 25 \%$  . Dengan perkataan lain validasi analisis multivariat yang diterapkan pada kasus data-data tersebut dianggap memuaskan.

#### 4. Kesimpulan

Dari hasil pembahasan di butir 3, terdapat tiga prediktor yang pengaruhnya sangat signifikan terhadap kondisi kualitatif geomagnet yaitu bilangan bintik matahari dalam orde harian di Lokasi SPM 2 , indeks  $\sum Kp$  di Lokasi SG 3 dan indeks  $\sum Kp$  di Lokasi SG 1. Selain itu penerapan analisis multivariat pada kasus data-data dalam makalah ini menunjukkan bahwa kondisi kualitatif geomagnet di lokasi atau sekitar lokasi-lokasi tersebut *tidak terganggu*. Untuk variabel bebas yang sudah dibuat menjadi baku (*standard*) dapat disimpulkan bahwa semakin besar nilai koefisien suatu variabel bebas semakin besar pula kuasa pemisah (*discriminant power*) dari variabel yang bersangkutan.

## 5. Daftar Pustaka

- Dillon, W.R. and Goldstein, M., (1984), *Multivariate Analysis: Methods and Applications*. Wiley Edition, New York.
- Lawson, M.P. and Cerveny, R.S., (1985), Seasonal temperature forecast as products of antecedent linear and Spatial temperature arrays. *Journal of Climate and Applied Meteorology*, 24 pp.848 – 859.
- Lehmiller, G.S. et al., (1997), Seasonal prediction models for North Atlantic basin hurricane location. *Monthly Weather Review*, 125, pp. 1780 – 1791.
- Maspupu, J., (2009), Analisis klaster pada data suimulasi indeks geomagnet lokal. *Prosiding Seminar Nasional Statistika IX, FMIPA Jur. Statistika ITS*, Surabaya, hal. 1 -11.
- Maspupu, J., (2011), Analisis Diskriminan pada data suimulasi indeks  $\sum K$ . *Prosiding Seminar Nasional Sains dan Pendidikan sains VI, FSM - UKSW*, Salatiga, hal. 847 -856.
- Miller, R.G., (1962), Statistical prediction by discriminant analysis. *Meteorology Monographs* 4, no. 25, *American Meteorological Society*, 53 pp.
- Tabachnick, B.G. and Fidell, L.S., (2007), *Using Multivariate Statistics*. Pearson International Edition, Boston.
- Ward, M.N. and Folland, C.K., (1991), Prediction of seasonal rainfall in the north Nordeste of Brazil using eigenvectors of sea-surface temperature. *International Journal of Climatology*, 11, pp.711 – 743.