

PRONOSTICO DE LA OCURRENCIA DE DESTELLOS MEDIANTE LA APLICACION DEL ANALISIS DISCRIMINANTE

Isabel Ferro Ramos, Jorge Pérez Doval
Instituto de Geofísica y Astronomía
Academia de Ciencias de Cuba

RESUMEN

Sobre la base de los datos de manchas solares obtenidos en el Departamento de Astronomía del IGA, se aplicó a los mismos el método del análisis discriminante para dos poblaciones, luego de haber dividido los datos según clases de Zurich y a su vez en dos poblaciones cada uno de ellos: que produjeron destellos a ± 1 día de considerados los datos de las manchas, y que no produjeron destellos.

Se obtuvo que la variable "área del grupo" es en general la más contributiva de forma independiente a la discriminación. Además, la variable "fase del ciclo" se hace importante a medida que se pase a clases de Zurich más complejas.

Para todas las clases de Zurich analizadas, ocurre que se pronostica mejor la no ocurrencia que la ocurrencia de destellos. La primera se logra pronosticar con efectividades entre 77% y 83%. En particular, el grupo de Zurich que mejor se pronostica es el E y el más difícil de pronosticar es el A. Los resultados logrados, utilizando un menor número de variables, son en cierta medida mejores que los reportados en otro trabajo similar, debido probablemente a haber utilizado entre las variables, la fase del ciclo y a haber subdividido el análisis según clases de Zurich.

INTRODUCCION

Las variables relacionadas con los grupos de manchas se han utilizado frecuentemente pero de manera unilateral, para relacionarlas con los destellos solares, con fines de pronóstico. (Antalova 1967, Kassinsky y Pobedin 1970, Knoska y Krivsky 1981). En particular esto concierne al área del grupo, la clasificación de Zurich, y el número de Wolf. Sin embargo, métodos multivariados que permitan tomarlas a todas en cuenta, se han utilizado menos, aunque en este sentido se deben mencionar los trabajos hechos empleando regresión multivariada (Jakimiec y Vasiucionek 1980) y otros basados en reconocimientos supervisado (Bernshtein y Burov 1976).

Sin embargo, se ha utilizado poco el método estadístico de reconocimiento de patrones, conocido por análisis discriminante (Hirman et al 1980). En el presente trabajo

se hará uso del mencionado método con dos particularidades no tomadas antes en cuenta; 1) la aplicación del método de forma separada a grupos caracterizados por una clase de Zurich dada y 2) la inclusión como variable de la fase del ciclo de actividad solar.

MATERIALES Y METODOS

Los datos de manchas fueron tomados del Servicio Optico Solar del Departamento de Astronomía del IGA correspondientes al período Marzo del 72 a Marzo del 78. Como se indicó, la muestra original fue dividida en 6 submuestras, correspondiendo cada cual a una clase de Zurich que se reseñan en la Tabla 1. Cada submuestra fue dividida en dos grupos: 1) el asociado al menos a un destello alguno en o al día siguiente de referirse los parámetros de las manchas y 2) el grupo de manchas no asociado a destello alguno en o al día siguiente en que corresponden los parámetros de las manchas.

Cada submuestra se analizó por separado, y al método se le entregó en cada caso una matriz de observaciones donde las filas son las manchas y las columnas sus variables correspondientes. Las variables utilizadas se muestran en la tabla 1.

La designación de ellas es la siguiente:

Sp-área del grupo de manchas en m_hv (millonésimas de hemisferio visible); Sp-variación respecto al día anterior en m_hs; # - número de Wolf para el grupo; del # - variación respecto al día anterior.

c - fase del ciclo de actividad codificado; Sm- área de la mancha mayor del grupo en m_hs, del Sm - variación respecto al día anterior de Sm; g - población del grupo que se codificó.

f - penumbra de la mancha mayor, que se codificó

CH - clasificación magnética, que se codificó:

El análisis discriminante es un método de reconocimiento de patrones que reduce un problema multivariado a uno univariado, mediante la construcción de la llamada función

discriminante, lineal en nuestro caso. Esta se construye como aquella combinación lineal de las variables que permiten producir la máxima diferencia entre grupos previamente definidos (en nuestro caso: dos). Tal función permitirá adjudicar cualquier nuevo individuo a uno de los grupos originales al evaluarlo según la variable que posea ese nuevo individuo (Davis 1978).

DESARROLLO

1. Posibilidad de la discriminación

Para cada submuestra se evaluó el estadígrafo F:

$$F = \frac{na + nb - m - 1}{(na + nb - 2)m} \cdot \frac{na \cdot nb}{na + nb} \cdot D^2$$

que se distribuye como la F de Fisher con m y $na + nb - m - 1$ grados de libertad, respectivamente, siendo m el número de variables y D la distancia de Mahalanobis.

El estadígrafo F responde a la hipótesis nula de que las dos medias multivariadas son iguales o que la distancia entre ellas es cero. Esto permite decidir si la discriminación es posible. Los resultados de aplicar la dócima a cada submuestra arrojan que dadas las variables utilizadas, la discriminación es siempre posible pues en todos los casos el estadígrafo supera el valor crítico para el nivel de significación de hasta 1%.

CONTRIBUCION DE LAS VARIABLES

En la propia tabla 1 se han subrayado las variables que (suponiendo sus efectos independientes) más contribuyen a la discriminación. En los grupos poco desarrollados de Zurich (A y B) ellas son el área de las manchas y su variación pero a medida que el grupo se vuelve complejo en su evolución según Zurich, la fase del ciclo se hace importante, con excepción de los grandes grupos E en que la población y el carácter de la penumbra de la mancha mayor, dominan la discriminación. En general, la variable más contributiva es el área del grupo, resultado que es de esperar.

EFFECTIVIDAD DE LAS CLASIFICACIONES

Los porcentos de mala clasificación se obtienen evaluando para cada observación, la función discriminante lineal, que se plotea por medio de barras a lo largo de la función discriminante. Sobre este eje también se plotean las puntuaciones discriminantes correspondientes a los centros de los dos grupos: R_a y R_b , así como R_o : el centro de los centros de los grupos A y B.

Los porcentos de mala clasificación del grupo A se definen como los individuos situados más allá de R_o hacia

el lado que ocupa el grupo B. Análogamente se define este porcentaje para el grupo B. Para cada submuestra se realizan varias corridas, involucrando diferente cantidad de variables. En general, los porcentos de mala clasificación aumentan cuando se disminuye el número de variables en cada submuestra, resultado obvio.

Por otra parte los resultados de los porcentos de mala clasificación para los conjuntos involucrados de variables se dan en la tabla 2. De esto se infiere que se pronostica mejor la no ocurrencia que la ocurrencia de destellos, hecho lógico pues el destello es un fenómeno más bien improbable que probable. El mejor grupo de Zurich que se pronostica es el E y el más difícil, el A.

Las clases de Zurich para las que se pronostica mejor la no ocurrencia de destellos son la B y la C y la más difícil la E. Por lo contrario, las clases de Zurich para las que se pronostica mejor la ocurrencia de destellos son la D y la E. Para cada clase, se calculó la efectividad E del pronóstico como la suma de los aciertos en las predicciones de ocurrencia y no ocurrencia entre el número total de predicciones. Ellas se dan en la propia Tabla 2.

El caso donde se obtiene la más alta efectividad (80%) es para la clase E. Para cada clase de Zurich se prepararon tablas de contingencia en que las entradas horizontales fue lo observado y las verticales lo pronosticado y se evaluó la verosimilitud del pronóstico ciego para cada una de ellas, así como la verosimilitud del método. Se obtuvo que para un nivel de significación de 1%, el método sugerido por la puntuación discriminante tiene derecho a existir, con excepción para la clase E en que la verosimilitud del método no es significativa. Esto se explica pues los grupos de manchas de clase E por lo general producen destellos, y se precisa de un método muy bueno para discriminar cuando no los van a producir. El método discriminante en las variables utilizadas, falla para esto.

La comparación con un trabajo anterior (Hirman et al 80) donde se emplea este mismo método pero se utiliza gran número de variables (29) arroja, en el caso de nuestro trabajo, porcentos de precisiones de ocurrencia que resultan ser similares (aproximadamente de 60%) o mayores (mayor que 70%) a las del trabajo mencionado como puede observarse de la Tabla 3. Este hecho de obtener parecidos resultados con un menor número de variables debe atribuirse a haber particionado previamente la muestra en clases de Zurich y haber tomado como variable la fase del ciclo.

CONCLUSIONES

1.- Para cada grupo de Zurich, las variables consideradas

.....

permiten siempre discriminar mediante una función discriminante lineal si el grupo producirá o no destellos con efectividades que se deducen de la Tabla 2.

2.- En general, de entre las variables consideradas, la más contributiva a la discriminación es el área del grupo de manchas. La variable "fase del ciclo" se hace importante para las clases de Zurich a partir de la clase C con excepción de la clase muy desarrollada E, en que la población del grupo y el carácter de la penumbra se hacen importantes.

3.- La efectividad más alta alcanzada en el pronóstico es para la clase E (80%) y la más baja para la clase A (69%). Sin embargo, el método no tiene derecho a existir para la clase E pues es el único caso en que la verosimilitud está por debajo de la verosimilitud del pronóstico ciego.

4.- Los porcentos de precisiones correctas o aciertos en las predicciones de ocurrencia son iguales o superiores a los del trabajo anterior de Hirman et al 1980. Este mejor resultado (utilizando un menor número de variables), debe estar relacionado con la partición de la muestra según clases de Zurich, y con haber tomado como variable, la fase del ciclo.

BIBLIOGRAFIA

Antalova, A. BAC No. 18, 2, 1967, pp 61-67

Bernstein, P.B., Burov, V.A. En "Surgimiento y evolución de las regiones activas ", Nauka, 1976, 186-189.

Davis, J. C. "Statistics and Data Analysis in Geology", N.Y., Wiley, 1973, 563 pp.

Hirman, J. W. et al. En "Solar Terrestrial Predictions Proceedings", vol. III, 1980, pp. C64-C74.

Jakimiec, M. Vasiucionek, J. En "Solar Terrestrial Predictions Proceedings", vol. III, 1980, pp. C54 - C63.

Kassinsky, V.V., Pobedin, L. F. Datos Solares, No. 9, 1970, pp. 106 - 110.

Knoska, S., Krivsky, L. BAC No. 32 - 5, 1981, pp. 292 - 303.

Tabla 1

Variables Importantes y tamaños de muestra

Submuestra	Clase de Zurich	Variables	N Muestra Grup. NDES	Grup. DES
I	A	Sp. del Sp, #, del #, c	58	33
II	B	Sp, del Sp, #, del #, c, SM, del SM,g	53	28
III	C	Sp, del Sp, #, del #, c, SM, del SM, g, f, CH	161	120
IV	D	Sp, del Sp, #, del #, c, SM, del SM, g, f, CH	176	176
V	E	Sp, del Sp, #, del #, c, SM, del SM, g, f, CH	22	114
VI	H,J	Sp, del Sp, #, del #, c, f, CH	96	127

Tabla 2

Porcientos de mala clasificación y efectividad del pronóstico

Clase de Zurich	NDES	Grupos DES	E
A	22%	45%	69%
B	17	43	74
C	19	33	75
D	26	38	68
E	32	18	80
H,J	23	35	70

Tabla 3

Porcientos de precisiones correctas o aciertos en las predicciones de ocurrencia

Clase de Zurich	%
A	58
B	64
C	72
D	71
E	93
H, J	79



COMISION FEDERAL DE ELECTRICIDAD

REVISTA MEXICANA DE GEOTERMIA

CONTENIDO

(Vol. 9 No. 3, Septiembre-Diciembre de 1993)

EDITORIAL	283
PERFORACION EXPLORATORIA EN LA ZONA DE TRES VIRGENES: EFECTOS POTENCIALES SOBRE EL AMBIENTE Alfredo Ortega, Cerafina Argüelles, et al.	285
ESTIMACION DE TEMPERATURAS DE LOS FLUIDOS DE PERFORACION DURANTE LA CIRCULACION Isaias Hernandez Ramirez, Alfonso Garcia Gutierrez y J. M. Morales Rosas	305
THE EVALUATION OF STEAM-WATER MIXTURES WITH SHARP EDGED CIRCULAR ORIFICES Pedro Sánchez Upton	321
PROGRAMA DE ESPECIALIDAD EN GEOTERMIA Gisela Montero Alpirez	345
INFLUENCIA DE LA CONDUCTIVIDAD TERMICA DE LA ROCA SOBRE LAS CONDICIONES TERMODINAMICAS INICIALES DE SISTEMAS HIDROTERMALES EN ROCA VOLCANICA Mano César Suarez Armaga	357
MODELO CONCEPTUAL APROXIMADO DEL YACIMIENTO GEOTERMICO EN EL VOLCAN DE EL CEBORUCO Ing. J. Cesar Viggiano Guerra	369
"WORLDWIDE UTILIZATION OF GEOTHERMAL ENERGY"	389
INTERNATIONAL SYMPOSIUM GEOTHERMICS 94 IN EUROPE	397
BIOGRAFIAS CONDENSADAS DE ALGUNOS AUTORES	403
PROGRAMA DE VOLCANOLOGIA Wolfgang E. Elston	408

La Revista "Geotermia" es un órgano de información técnica publicada por la Gerencia de Proyectos Geotermoeléctricos de la Comisión Federal de Electricidad. Edición cuatrimestral y distribución gratuita. Certificados de Licitud de Título No. 2784, Octubre 2, 1985; Licitud de Contenido No. 2514, Diciembre 1, 1986, expedidos por la Comisión Calificadora de Publicaciones y Revistas Ilustradas de la Secretaría de Gobernación. Correspondencia de 3da. Clase, Registro LXC 046-0488. Características 210122112. GEOTERMIA-LA REVISTA MEXICANA DE GEOENERGIA es nombre registrado en la Dirección General de Derechos de Autor de la SEP. Reserva 01 86. Reservados todos los derechos de reproducción del material publicado. Registro ISSN 0186-5897. Miembro Corporativo de la Asociación Geotérmica Internacional (IGA) y de la Asociación Geotérmica Mexicana (AGM).