Geofísica Internacional (1992), Vol. 31, Num. 2, pp. 151-153

Expert systems for forecasting solar flares associated to C and D Zurich classes

I. Ferro Ramos, J. Pérez Doval and T. Cepero Chao
Institute of Geophysics and Astronomy
Cuban Academy of Sciences, Havana, Cuba

RESUMEN

Con la ayuda del medio ambiente ARIES fueron obtenidos dos sistemas expertos para el pronóstico de la ocurrencia de destellos asociados a las clases de Zurich C y D, uno para el pronóstico diario de la situación destellante (sistema experto FUL) y otro para pronosticar los destellos que ocurran dentro de las 24 horas después de la hora de observación del grupo de manchas asociado (sistema experto FULGOR). La efectividad general de las bases de conocimiento es 80% . Ellas operan sobre el 44% de las manchas-día, las cuales producen el 66% de los destellos. En ambos casos, la calidad del pronóstico es superior a la efectividad del mejor de los pronósticos triviales, el pronóstico ciego.

PALABRAS CLAVE: Destellos solares, pronóstico solar.

ABSTRACT

With the help of shell ARIES two expert systems are obtained for forecasting the occurrence of flares associated with the C and D Zurich classes: one for daily forecasting of the flare situation (expert system FUL) and the other for flares that occur 24 hours after the observation of the associated sunspot group (expert system FULGOR). General effectiveness of the bases is 80%. They operate upon 44% of sunspot-day which produce 66% of flares.

In both cases, the quality of the forecast is better than the effectiveness of the best trivial forecasting, the blind one.

KEY WORDS: Solar flares, solar forecasting.

INTRODUCTION

The forecasting of solar flares (Ferro, 1989) belongs to the forecasts of solar activity of short and supershort term type. They are based upon the characteristics of the active region: morphology and magnetic class of sunspots, their quantitative physical parameters, topology of filaments, brightness and compactness of flocculi, situation relative to active longitudes, precedent flare activity, reinforcement of the X and UV emission, intensification of coronal lines, etc.

For the forecasting of flares two types of methods have been used: the synoptic and the mathematical. The former has a strong subjective character while the latter has usually been used throughout by means of univariate and multivariate regression equations, graph theory and pattern recognition techniques. All provide an effectiveness not greater than 86-88% (McIntosh, 1978).

With the appearance of artificial intelligence techniques, the mathematical method has been enriched with a fourth branch: expert systems based upon knowledge. The first objective of this paper is to create an expert system for forecasting the occurrence of at least one chromospheric flare of any type on the day of observation in accordance with the mean characteristics of sunspots of that day and the preceding one. The second expert system gives information if the active center will produce at least one flare 24 hours after the observation of the associated sunspot group.

For that purpose, we employed the shell ARIES, created in Cuba under the guidance of Dr. Julio J. Valdés Ramos and Lic. Argelio V. de la Cruz (Valdés and De la Cruz, 1988). The bases of knowledge were created with the help of four artificial intelligence algorithms implemented through the module SAMC (Knowledge Acquisition and Handling System).

DATA AND METHODS

Source data were the observations of sunspots carried out at the Institute of Geophysics and Astronomy of the Cuban Academy of Sciences from March 1972 to March 1978. The parameters employed are: (1) Sunspot area, (2) Variation of sunspot area in relation to the previous day, (3) Number of pores, (4) Variation of the number of pores in relation to the previous day, (5) Phase of the cycle divided in three categories: before, at and after minimum (1975 and 1976 were considered as minimum years), (6) Area of the largest sunspot, (7) Variation of the largest sunspot area, (8) Population of the group classified into (a) open, (b) intermediate, (c) compact. Each of the categories is subdivided into three groups according to whether they have the leader spot bigger, the next spot bigger, or both sunspots equal size. (9) Morphology of the penumbra; the McIntosh classification was used that considers 5 types: (a) rudimentary (b) small and symmetric, (c) big and symmetric, (d) small and asymmetric, (e) big and asymmetric. (10) The magnetic class. The Mount Wilson classification was used and 5 types were considered: alpha, beta, beta-gamma, gamma and delta. The categorization of variables was different in accordance to the Zurich class of sunspots.

Source data for flares were taken from the Quarterly Bulletin on Solar Activity.

For the forecasting of the solar flare situation we considered a sunspot to be associated to a flare when at least one flare had occurred on the day of observation of the sunspot. With the purpose to clearly define the non-flare situation we took as "sunspot not associated to flares" any case where the last flare had occurred 48 hours before or after the observation.

For the second objective of this paper, i.e. to achieve a forecasting of the flare occurrence after the time of observation of a sunspot group, but within that day, we adopted as training matrices those sunspot-day groups that produced flares after and before or after the time of observation.

The methods employed are four automatic learning algorithms called ALV. They are: asymmetric lambda, symmetric lambda, error reduction and one based upon entropy (Gil, 1989).

DISCUSSION AND RESULTS

Several knowledge bases with different categorizations of the variables were created with the help of program LEARN of module SAMC.

The general base of data (901 sunspots-day) was divided according to Zurich classes and to each class was applied a particular categorization of variables.

For the bases corresponding to Zurich classes A, B, E and H we generated exact rules with the aid of four algorithms and the condition was imposed that they be filled by at least three individuals. We considered rules whose probability values were larger that 70%.

For the bases corresponding to Zurich classes C and D we used the same criteria but only the lambda asymmetric algorithm was used.

VALIDITY OF FORECASTS

Following Knoska and Krivsky (1981), the frequency of flares according to sunspots is stated in Table 2, and the frequency of sunspot-day is stated in Table 3 (Doval, 1990).

Table 1

Zurich class		Frequency
A	-	10%
B	-	15%
C	-	26%
D	-	40%
E	-	65%
H	-	19%

Table 2

Zurich class		Frequency
A	-	14%
B	-	9%
C	-	20%
D	-	24%
E	-	7%
H	-	26%

According to these tables for 100 sunspot-day, we obtain:

Table 3

Type	Produced flares	
	Yes	No
A	1	13
B	1	8
C	5	15
D	10	14
E	5	2
H	5	21

In accordance with these proportions 200 sunspot-days were chosen that produced at least one flare-day or none at all. The results of the verification were arranged in contingency tables of 2x2 (Table 4).

Table 4

observed	YES	NO	
Forecast			
YES	C11	C12	N1
NO	C21	C22	N2
	N1'	N2'	N

We now calculate for each Zurich class, the likelihood of the forecast, defined as:

$$P = \frac{C11 + C22 - C21 - C12}{N}$$

On the other hand, the best of trivial forecasts (the blind forecast) has a likelihood

$$P' = \frac{|N2 - N1|}{N}$$

It is necessary to take into account the likelihood of the forecast because a forecasting method is good only if it is significantly better than a trivial forecast. The effectiveness was calculated as:

$$E = \frac{C11 + C22}{N}$$

and the rates of assessments when the forecast is YES and when it is NO, as:

$$E_{YES} = \frac{C11}{N1} \qquad E_{NO} = \frac{C22}{N2}$$

For both objectives it is found that the only sufficiently good forecasts are those for the Zurich classes C and D, because only they are superior to the blind forecast.

Also the effectiveness is the same for the forecast of the flare situation and for the forecast of flares after observation of the associated sunspot-group. In both cases it is 80% for the combined classes C and D.

Also for the two combined classes, the effectiveness obtained in the YES forecast is superior to that reported by Hirmann *et al* 1980 at the Boulder Forecasting Center A, for the forecast of X flares with the aid of traditional methods and using more than 100 variables.

It should be remarked that the two Zurich classes (C and D) that can be forecast are precisely those generated with a bigger learning matrix: consequently the base possesses more quantity of rules. Both Zurich classes constitute approximately half of the sunspot-day and this indicates the utility of the method.

BIBLIOGRAPHY

DOVAL, J. P., 1990. Personal communication.

FERRO RAMOS, I., 1989. Conferences about solar flares at the XVI International School for Youth Astronomers, Cuba.

HIRMAN, J.W. *et al.*, 1980. The application of multivariate discriminant analysis to solar flare forecasting. *In:* Solar Terrestrial Predictions Proceedings, Vol. III.

KNOSKA, S. and L. KRIVSKY, 1981. *Bol. Astronom. Inst. Chekoslovakia (BAC), 32,* 5, 292-303.

MICHAELSEN, R. et al. 1985. The Technology of experts systems, in Byte, April, pp. 303-312.

VALDES RAMOS, J., A. DE LA CRUZ RIVERO, 1988. ARIES, a project of artificial intelligence for the development of expert systems. ICIMAF, ACC, 44 pp.

McINTOSH, P., 1978. Observations and forecast of solar activity, Gil Chaviano, M. Licence Thesis Faculty of Mathematics, Havana University, 1989.

I. Ferro Ramos, J. Pérez Doval and T. Cepero Chao

Institute of Geophysics and Astronomy, Cuban Academy of Sciences,
Av. 212 No. 2906
La Lisa, Cd. de la Habana, Cuba