Journal of The Korean Astronomical Society
40: 99 ~ 106, 2007

PREDICTION OF DAILY MAXIMUM X-RAY FLUX USING MULTILINEAR REGRESSION AND AUTOREGRESSIVE TIME-SERIES METHODS

J.-Y. Lee[1], Y.-J. Moon[1], K.-S. Kim[1], Y.-D. Park[2] and A. B. Fletcher[2*]
[1]Department of Astronomy and Space Science, Kyung Hee University, Yongin, Gyeonggi 446-701, Korea
e-mail: jlee@khu.ac.kr & *moonyj@khu.ac.kr*
[2]Korea Astronomy and Space Science Institute, 61-1 Whaam-dong, Yuseong-gu, Daejeon 305-348, Korea

ABSTRACT

Statistical analyses were performed to investigate the relative success and accuracy of daily maximum X-ray flux (MXF) predictions, using both multilinear regression and autoregressive time-series prediction methods. As input data for this work, we used 14 solar activity parameters recorded over the prior 2 year period (1989–1990) during the solar maximum of cycle 22. We applied the multilinear regression method to the following three groups: all 14 variables (G1), the 2 so-called 'cause' variables (sunspot complexity and sunspot group area) showing the highest correlations with MXF (G2), and the 2 'effect' variables (previous day MXF and the number of flares stronger than C4 class) showing the highest correlations with MXF (G3). For the advanced three days forecast, we applied the autoregressive time-series method to the MXF data (GT). We compared the statistical results of these groups for 1991 data, using several statistical measures obtained from a 2 × 2 contingency table for forecasted versus observed events. As a result, we found that the statistical results of G1 and G3 are nearly the same each other and the 'effect' variables (G3) are more reliable predictors than the 'cause' variables. It is also found that while the statistical results of GT are a little worse than those of G1 for relatively weak flares, they are comparable to each other for strong flares. In general, all statistical measures show good predictions from all groups, provided that the flares are weaker than about M5 class; stronger flares rapidly become difficult to predict well, which is probably due to statistical inaccuracies arising from their rarity. Our statistical results of all flares except for the X-class flares were confirmed by Yates' χ^2 statistical significance tests, at the 99% confidence level. Based on our model testing, we recommend a practical strategy for solar X-ray flare predictions.

Key words : Sun: solar activity — Sun: X-ray flares — Sun: flare prediction — Sun: statistical method

I. INTRODUCTION

A solar flare is the solar atmospheric response to a sudden release of energy associated with magnetic reconnection accompanying an acceleration of electrons, protons and heavier ions. Accurate solar flare prediction becomes increasingly important, nowadays, as these high energy particles can seriously affect satellite operations, communications and navigation, as well as electric power distribution networks. To fulfill this need, solar X-ray flare predictions have been studied using several methods. These statistical analyses have been used to find the effective contributions of each of the observed solar activity variables to solar flare phenomena.

Over the past decade, several studies attempted to predict the daily total and maximum X-ray (1–8Å) flare flux values, using the multilinear regression of multivariate vectors containing more than 10 solar activity variables. These usually included sunspot group and X-ray flare information. Bornmann and Shaw (1994) found a good relationship between the daily mean flaring rate of X-ray flares and each of the McIntosh classification parameters, by using a classical multilinear regression method. Also, several statistical methods, such as robust regression (Bartkowiak and Jakimiec, 1990a; Bartkowiak and Jakimiec, 1990b), two-step regression (Jakimiec, 1993) and distance-based regression (Jakimiec and Bartkowiak, 1994), have been applied to these variables, in a search for more reliable predictive flare models, but they do not show any remarkable improvement over the classical multilinear regression methods. We also note that previous studies have not presented any systematic evaluations of the statistical significance of their predictions.

To develop the flare prediction model, the detailed analyses of the occurrence of X-ray flare and the physical properties of an active region have been studied more specifically by several authors. Wheatland (2000) presented that the waiting time distribution of flares is consistent with a time-dependent Poisson process and the probability of flare occurrence per unit time follows an approximate exponential distribution based on 25 years of soft X-ray flare. Moon et al. (2001) showed that the waiting time distribution of flare follows a Poisson probability function and a 2-

Corresponding Author: Y.-J. Moon
*Present Address: School of Physics, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia

3 day period is best to provide the mean flaring rate to be used in flare prediction. A flare prediction system based on the Poisson statistics using the McIntosh classification parameters has been developed by Gallagher et al. (2002). These flare prediction models presented the probabilities of the flare production. On the other hand, Barnes and Leka (2006) showed that the coronal field topology derived from a magnetic charge topology (MCT) model provides some useful information for distinguishing the flaring and flare-quiet region. Most recently, the studies of the major solar flare-forecasting showed that the effective connected magnetic field (Georgoulis and Rust, 2007) and the unsigned flux near compact, high-gradient polarity-separation lines (Schrijver, 2007) could be efficient criteria for flare forecasting.

In this paper, we make a statistical study of daily maximum X-ray flux predictions, using a classical multilinear regression method. In addition, an autoregressive time-series method is employed to forecast the MXF in the advanced three days. The methods are applied to 14 recorded solar activity variables from 1989 to 1990 to predict the daily MXF. Using these results, we predict the daily MXF for 1991 data and evaluate the statistical results of the prediction using contingency table and Yates' χ^2 test. The main objectives of our study are to examine which of these input data are useful for making reliable solar flare predictions, and to estimate how accurately such predictions can be made. Finally, we suggest a simple, daily, X-ray flare forecasting model which depends on just 2 solar activity variables; this simple model shows the best predictive performance among the set of models considered here.

In §II we describe 14 variables used in our prediction models. In §III we present the multilinear regression and autoregressive time-series prediction methods. In §IV we describe the evaluation methods, the contingency table and Yates' χ^2 test. In §V the statistical results are evaluated and compared for the 4 groups. In §VI we summarize the several results and present the simple model for the MXF forecasting.

II. DATA

During a solar maximum period, the prediction of X-ray flares gives relatively good results when the solar activity data are used within a suitable statistical model (Lee *et al.*, 1999). In this paper, we tried to estimate the daily maximum X-ray flare flux values (MXF) for 1991, using data recorded over the prior 2 year period (1989–1990) during the solar maximum of cycle 22. These solar activity data were obtained primarily from the Solar and Upper Atmospheric Data Services website (http://www.ngdc.noaa.gov/stp/SOLAR/solar.html) of the U.S. National Geophysical Data Center (NGDC). In a few cases, data from the NGDC Space Physics Interactive Data Resource (SPIDR: http://spidr.ngdc.noaa.gov/spidr/) were used.

Table 1 presents the 14 kinds of solar activity variables used in our predictive models. These variables describe the daily characteristics of solar activity, specifically, the sunspot group and X-ray flare information. In this study, we classified these data into two types, according to whether the variable is a 'cause' or an 'effect' of a flare event. If it is a potential cause of flares, it is categorized as a cause variable, otherwise it is an effect variable. The sunspot group information is provided by 6 variables: the 3 McIntosh Parameters (the modified Zurich class Z, p and c; abbreviated here as ZM, PM and CM, respectively); the number of Sunspots (SS); the Sunspot Group Area (SA); and the Magnetic Class (MC). We collectively named these 6 quantities the 'cause' variables. The remaining 8 'effect' variables give the flare information: the daily Maximum (MXF) and Average X-ray Fluxes (AXF); the 2800 MHz Radio Flux (RF); the number of C ($10^{-6} \leq I(W/m^2) < 10^{-5}$); M ($10^{-5} \leq I < 10^{-4}$); and X ($I \geq 10^{-4}$) Class Flares (CF, MF and XF, respectively); the number of strongs flares (SF) stronger than C4 ($4 \times 10^{-6}\, W/m^2$); and the number of faint flares (FF) stronger than C2 ($2 \times 10^{-6}\, W/m^2$) but weaker than C4.

The 3 McIntosh parameters (McIntosh, 1990), which give information on the sunspot groups, are determined by observations of, the length of the sunspot group (Z), the size and shape of the penumbra of the largest sunspot in the group (p), and the distribution of spots within the group (c). The statistical correlation of the McIntosh parameters with observations of X-ray flaring has been investigated (e.g., Bornmann, Kalmbach, and Kulhanek, 1994), and these (Z, p, c) parameters have been used for solar activity predictions (Neidig *et al.*, 1986; Bartkowiak and Jakimiec, 1986; Jakimiec and Bartkowiak, 1986; Jakimiec, 1993 Lee and Kim, 1996). Following Jakimiec (1993), we analyzed only those sunspot groups with Z McIntosh parameter (otherwise known as the modified Zurich class) equal to D, E or F; only bipolar sunspot groups, with penumbrae on both ends of each group, were thus included. We also took care to make separate statistical analyses of the data, for each of the 'growing' and 'decaying' phases of sunspot group evolution; those groups in the 'growing' phase present the areas which increase with time, while those in the 'decaying' phase present the areas which decrease with time – which results in somewhat different flare productivity. All McIntosh parameters were converted into integers to enable further numerical calculations (see Neidig *et al.*, 1986, for details).

III. MODELS AND ANALYSIS

We first examined the statistical correlations between the maximum X-ray flare flux (MXF; on the prediction day) with each of the 14 solar activity varaibles one day earlier. The correlation coefficients are listed in Table 1. These coefficients were used to judge which variables might be useful for X-ray flare predictions, by considering the strength of their correlations with

TABLE 1.
CORRELATION OF VARIOUS SOLAR ACTIVITY VARIABLES WITH MAXIMUM X-RAY FLUX

SOLAR ACTIVITY VARIABLE	ABBRV	CAUSE OR EFFECT	CORRELATION VALUE GROWING* PHASE	DECAYING* PHASE	GROUP MEMBERSHIP
Z McINTOSH PARAMETER	ZM	C	0.16	0.28	G1
P McINTOSH PARAMETER	PM	C	0.15	0.13	G1
C McINTOSH PARAMETER	CM	C	0.42	0.34	G1, G2
NO. OF SUNSPOTS	SS	C	0.33	0.31	G1
$\log_{10}$(SUNSPOT GROUP AREA)	SA	C	0.32	0.40	G1, G2
$\log_{10}$(MAXIMUM X-RAY FLUX)	MXF	E	0.50	0.45	G1, G3, GT
$\log_{10}$(AVERAGE X-RAY FLUX)	AXF	E	0.47	0.48	G1
NO. OF STRONG FLARES	SF	E	0.53	0.43	G1, G3
NO. OF FAINT FLARES	FF	E	0.39	0.44	G1
NO. OF C FLARES	CF	E	0.21	0.01	G1
NO. OF M FLARES	MF	E	0.52	0.42	G1
NO. OF X FLARES × 10	XF	E	0.31	0.27	G1
MAGNETIC CLASS	MC	C	0.16	-0.02	G1
$\log_{10}$(2800 MHz RADIO FLUX)	RF	E	0.46	0.39	G1

* : See section II

MXF; possible predictive groups of variables were then chosen, as detailed in the following sections.

(a) Groups G1–G3

For the purposes of daily maximum X-ray (MXF) flux prediction, we defined 4 groups (G1 through GT) of solar activity variables. The first 3 groups were defined as follows. We included all 14 variables for G1. For each of G2 and G3, we chose the 2 variables showing the highest correlations of MXF with the 'cause', and 'effect', variables, respectively. For G2, this meant the c McIntosh parameter (CM) and the sunspot group area (SA). For G3, this meant the maximum X-ray flux (MXF) and the number of 'strong' flares stronger than C4 (SF). The prediction variables from each of these 3 groups were substituted into Equation (1) below, which is a linear regression in multiple variables (x_{ij}):

$$y_{i+1}(X_i) = a_0 + \sum_{j=1}^{n} a_j x_{ij}, \qquad (1)$$

where the input data vector is $X_i = [x_{i1}, x_{i2}, x_{i3}, \cdots, x_{in}]$. y_{i+1} is the MXF value calculated for the $(i+1)$-th prediction day. n is the number of solar activity variables chosen for making the group prediction ($n = 14$ for G1 and $n = 2$ for G2–G3) and x_{ij} is the value of the j-th solar activity variable of the group, as evaluated on the i-th day. Interactive Data Language (IDL) programs were used to calculate the regression coefficients a_0 and a_j for solar data recorded over a period of two years (1989 Jan 1–1990 Dec 31). To predict the MXF values for each day $(i+1)$ in 1991, we use these predetermined coefficients, and the values (x_{ij}) of the solar activity variables recorded on the preceding days (i), in Equation (1).

(b) Groups GT

For group GT, we applied an entirely different analysis, based on a simple, autoregressive, time-series prediction method. Here, the input data were drawn solely from the 1991 time series of MXF values. We attempted to predict MXF for 3 consecutive days, the prediction day (t), as well as the following 2 days ($t+1$, $t+2$). Using IDL functions (ts_fcast.pro), a q-th order, autoregressive model was used to obtain the future MXF values (y_t, y_{t+1}, y_{t+2}), using equation

$$y_t = b_0 + b_1 y_{t-1} + b_2 y_{t-2} + \cdots + b_q y_{t-q}. \qquad (2)$$

The input data, for each of these 3 predictions, were the 1991 MXF values taken from the q days prior to each successive prediction day (for t, $t+1$ and $t+2$, the corresponding groups are GT.1, GT.2 and GT.3, respectively). For example, if q was chosen to be 2, then the GT.1 values used to predict y_t would be (y_{t-2}, y_{t-1}), for y_{t+1}, the GT.2 values would be (y_{t-1}, y_t), and for y_{t+2}, the GT.3 values would be (y_t, y_{t+1}). The same autoregression coefficients b_i were used for each successive forecast. The b_0 coefficient is unavoidably affected by uncorrelated random errors. All autoregressive predictions were carried out in the following two stages:

1. In the 1st stage, the optimum value of q was determined from trial predictions of MXF values for the 3 days prior to the initial prediction day t: (y_{t-3}, y_{t-2}, y_{t-1}). The autoregression coefficients b_i were determined by minimizing the uncorrelated random error terms b_0 in Equation (2). The number q, of prior days used as input data for the trial predictions, was varied from 2 through 7. The success of these trials was gauged from

forecasts of Event 1 (See section V) flares. The optimal q value was chosen to yield the largest sum of the 'Success Ratio' (SR) and the 'Probability of Detection' (POD) (See section IV). For the 11 days in 1991 for which there were neither observations, nor predictions, of Event 1 flares for any of the prior 3 days, we chose $q = 2$ by default. The histogram of the number of optimally selected q values is shown in Figure 1.

2. In the 2nd stage, the autoregressive prediction of the future 3-vector (y_t, y_{t+1}, y_{t+2}) was calculated using Equation (2) again, this time with the preselected optimum q value from the 1st stage trial predictions.

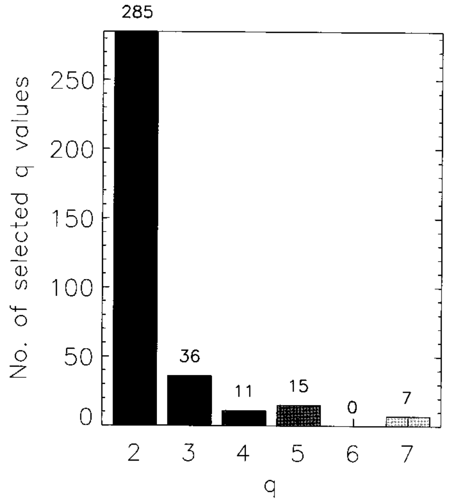

Fig. 1.— Histogram showing the number of optimally selected q values found from trial predictions for the 3 days prior to the initial prediction day, plotted against the number q of input days used in the autoregression defined in Equation (2). The number of optimal q values selected appears above the qth bar. Only 354 days out of 365 are shown here; the other 11 days in 1991 had no observed or predicted Event 1 flares (these days were assigned $q = 2$ by default).

IV. FORECAST EVALUATION METHODOLOGY

To assess the success of our statistical prediction models, we employed a simple 2×2 contingency table, which is commonly used for forecast evaluations at the Space Weather Prediction Center (*http://www.swpc.noaa.gov*). This tool is useful for understanding how the model forecasts compare to the observed flare events (see Table 2). In this paper, A is the number of observed flare events that were successfully forecasted, B is the number of forecasted flare events that proved to be false alarms, C is the number of observed flares that failed to be forecasted, and D is the number of correctly forecasted non-events. Table II is also useful for establishing the definitions of several statistical measures, namely: 'Success Ratio' (SR), 'Probability of Detection' (POD) and 'Percent Correct' (PC), which are defined as follows:

$$SR \equiv A \;/\; (A+B), \tag{3}$$
$$POD \equiv A \;/\; (A+C), \tag{4}$$
$$PC \equiv (A+D) \;/\; (A+B+C+D). \tag{5}$$

SR is thus a measure of how successful the predictions are, considering just the forecasted flare. POD is the probability of detecting observed flares via the chosen prediction method. PC is the percentage of correct predictions, counting both flare 'events' and 'non-events'.

TABLE 2.
CONTINGENCY TABLE

		Event Observed	
		Yes	No
Event Forecast	Yes	A	B
	No	C	D

We also evaluated the statistical significance of our forecasts, by applying a Yates' χ^2-test to each contingency table. This test assumes the null hypothesis that the rows and columns of the table are statistically independent (i.e. that the forecasted and observed flare events are totally uncorrelated). Yates' χ^2-test is a simple extension of the standard χ^2 goodness-of-fit test, and its main results are: the Yates' chi-squared test statistic (χ^2), the confidence level (CL; the percentage probability of obtaining a chi-squared value less than χ^2), and the coefficient of contingency (CC; a positive scalar in the interval [0.0, 1.0], which measures the degree of dependence within the contingency table). Yates' χ^2-test is usually employed to accept or reject the null hypothesis (independence) at the 99% (or 95%) confidence level; if $\chi^2 > 6.635$ (or 3.841), it is rejected with 99% (or 95%) confidence, implying that there does exist a statistical dependence between the row and column variables. Yates' method is regarded as being more appropriate for the testing of 2×2 contingency tables than the standard χ^2-test (Langley, 1971). Wilson (1994) showed that the standard χ^2-test overstates the confidence level for a 2×2 contingency table, relative to that from Yates' test.

V. RESULTS AND DISCUSSIONS

The comparative evaluation of forecasts is usually a complex process, depending on detailed, quantitative assessments of their quality. We tried to evaluate our statistical prediction models in various ways, using the statistical measures described in §IV. The resulting performance values are presented in Table 3. In the top section of this Table, we give the number of correct predictions (A+D) for all classes of solar X-ray flares, namely. This is to test the accuracy of our models for forecasting solar flare activity in general, irrespective of flare strength. A prediction of an observed flare was counted as correct if the base 10 logarithms of the predicted and observed MXF values differed by less than 0.5. We also defined 3 event categories: flares M class or stronger (Event1 or E1); flares M5 class or stronger (Event2 or E2); and X class flares (Event3 or E3). We defined E1 to test our forecasts for all of the 'strong' flares. Flares M5 class or stronger (E2 or E3) are regarded as those which significantly affect our space environment. The relatively weak C class flares are very common during solar maximum. Although long duration C class flares may affect the solar-terrestrial environment(Kahler *et al.*, 1984; Kahler, 1992; Hudson, Haisch, and Strong, 1995), we did not consider X-ray flare duration in this analysis.

In the following subsections, we compare the performances of our predictions by group, event, measure and confidence level. We also defined a new measure, the Overall Performance (OP) average:

$$OP \equiv (SR + POD + PC) / 3. \quad (6)$$

OP helped us judge which groups made the most reliable predictions, on considering the average over all 3 measures (SR, POD and PC).

(a) Groups: G1–G3

We first compare the results from each of G1–G3 and GT, separately, as they were analyzed using different prediction methods. We excluded G2 since the 'cause' variables (CM and SA) underpredict the flare fluxes; we can see this from the very low number of successfully forecasted A (see Table 2) for E2 and E3 (i.e. A = 1 and A = 0, respectively). Briefly speaking, for the multilinear regression predictions, the best predictor is G1 or G3 (See OP in Table 3). We note that the G3 is a reliable MXF predictor though it consists of only two 'effect' variables (SF and MXF) compared with 14 variables of G1.

(b) Groups: GT

For groups GT, we did not expect unreliable autoregressive predictions from the MXF time series, as MXF is not a 'cause' variable; among the 'effect' variables, the MXF naturally has the strongest correlation with its prior day value (see Table 1). The differences of the statistical measures between GT groups are statistically not so large, except for the 'extreme' E3 events (X class flares). The optimum value (q), 2 was evaluated (see Fig 1 and §III (b)) as the best number of actual time-series values to be used in the forecast. This implies that the x-ray flux for previous two days is an useful proxy for flare forecasting, which is consistent with Moon et al.(2001) who found from the waiting time distribution of flares that the period for a constant mean flaring rate is a 2-3 days.

(c) Events: E1–E3

We now compare the Overall Performance (OPn) averages between the various Events En, in more detail. For the E1 flares, the multilinear (G1 and G3) forecasts clearly perform better than the GT.n autoregressions. For the E2 flares, the GT.n autoregressive predictions together perform almost as well as the multilinear forecasts using G1 and G3. On the other hand, the situation is reversed for the E3 flares, for which the autoregressive forecasts are clearly more reliable. GT.3 is the best predictor to use for X flares. We note the drop in OPn performance, on going from the E1 flares to the E3 flares. While all statistical measures (SR, POD, PC and OP) exceed ≈60% for the E1 flares, this quickly ceases to be true for E2 and E3 flares, for which all measures fall below ≈50%, except for PC (it actually increases with flare strength, a trend which can be easily explained – see the next section (d)).

(d) Measures: SR, POD, PC and OP

We now look at how the performances of the various measures vary according to the flare strength. We expect two trends with increasing flare strength: the first is that since stronger observed flares are rarer than weaker ones, we expect (A+C) in Table 2 to decrease significantly and (B+D) to increase by the same amount as we consider increasingly strong flares. The second is that because strong predicted flares are rare, they should be harder to forecast reliably than the numerous weak flares, due to the statistical inaccuracies in the prediction methods used. Thus, we expect (A+B) to decrease significantly and (C+D) to increase by the same amount, as we consider stronger flares. Based on these 2 trends, we expect the Table 2 counts to have the limiting values (A = 0; B = 0; C = 0; D = N) for flares of arbitrarily high flux ($I \rightarrow \infty$). The 4 statistical measures should then have the limiting behaviour:

$$\lim_{I \rightarrow \infty} SR \rightarrow 0; POD \rightarrow 0; PC \rightarrow 1; OP \rightarrow 1/3. \quad (7)$$

These expectations are confirmed by an example G1 Table 3:
SRn% decreases severely from E1(75.2) to E3(23.1);
PODn% decreases severely from E1(90.0) to E3(6.3);
PCn% increases significantly from E1(75.1) to E3(84.9);
OPn% decreases severely from E1(80.1) to E3(38.1).

(e) Confidence Level and Coefficient of Contingency

We estimated the Yates' χ^2 test statistic, the confidence level (CL) and the coefficient of contingency (CC), for each group and event pair, by applying Yates' χ^2-test to each contingency table; the results are summarized in Table 4. It is found that the correlations between the observed and forecasted E1 and E2 flares are statistically significant at greater than the 99% confidence level ($\chi^2 > 6.635$). For E3 events, the dependence between the observed and forecasted flares is sta-

TABLE 3.
COMPARATIVE EVALUATIONS OF SOLAR MAXIMUM X-RAY FLARE FORECASTS BY GROUP, EVENT AND MEASURE

MEASURE	G1	G2	G3	GT.1	GT.2	GT.3
TOTAL NO. OF OBSERVATION DAYS A+B+C+D = 365						
NO. OF CORRECT PREDICTIONS	228	230	232	221	206	206
PERCENT CORRECT (PC%)	62.5	(63.0)	**63.6**	60.5	56.4	56.4
PERFORMANCE EVALUATIONS USING THE CONTINGENCY TABLE						
Event 1 : M FLARE OR STRONGER (NO. OF DAYS WITH OBSERVED EVENTS A+C = 229)						
A	206	208	199	179	166	157
D	68	59	69	78	73	70
SR1 (%)	*75.2*	(73.0)	*74.8*	**75.5**	72.5	70.4
POD1 (%)	**90.0**	(90.8)	86.9	78.2	72.5	68.6
PC1 (%)	**75.1**	(73.2)	73.4	70.4	65.5	62.2
OP1 (%)	**80.1**	(79.0)	78.4	74.7	70.2	67.1
Event 2 : M5 FLARE OR STRONGER (NO. OF DAYS WITH OBSERVED EVENTS A+C = 89)						
A	23	1	24	27	23	21
D	251	276	249	244	242	246
SR2 (%)	**47.9**	(100)	*47.1*	45.8	40.4	41.2
POD2 (%)	25.8	(1.1)	27.0	**30.3**	25.8	23.6
PC2 (%)	**75.1**	(75.9)	*74.8*	*74.3*	72.6	73.2
OP2 (%)	*49.6*	(59.0)	*49.6*	**50.1**	46.3	46.0
Event 3 : X FLARE (NO. OF DAYS WITH OBSERVED EVENTS A+C = 48)						
A	3	0	2	6	6	10
D	307	317	312	296	301	305
SR3 (%)	23.1	(*)	28.6	22.2	27.3	**45.5**
POD3 (%)	6.3	(0)	4.2	12.5	12.5	**20.8**
PC3 (%)	84.9	(86.9)	*86.0*	82.7	84.1	**86.3**
OP3 (%)	38.1	(@43.4)	39.6	39.2	41.3	**50.9**
All Events : (E1 + E2 + E3)/3						
SR (%)	48.7	(@86.5)	50.2	47.8	46.7	**52.3**
POD (%)	**40.7**	(30.7)	39.4	*40.3*	36.9	37.7
PC (%)	**78.4**	(78.6)	*78.1*	75.8	74.1	73.9
OP (%)	**55.9**	(@62.6)	**55.9**	54.7	52.6	54.6

Bold values are from the best performing groups.
Emphasized values are from the 'runner-up' groups (i.e. within 1% of the best groups).
( ) : The values come from the excluded model (i.e. from group G2).
* : undefined value for SR3% – this is excluded from calculations of averages.
@ : 1 or more statistical measures excluded in calculation of average.

TABLE 4
YATES' χ^2-TESTS FOR STATISTICALLY SIGNIFICANT CORRELATIONS

EVENT		G1	G2	G3	GT.1	GT.2	GT.3
E1	χ^2	**70.67**	**56.37**	**59.25**	**45.73**	**23.88**	**13.57**
	CL	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.98**
	CC	0.40	0.37	0.37	0.33	0.25	0.19
E2	χ^2	**15.16**	0.36	**15.13**	**16.09**	**8.34**	**8.04**
	CL	**99.99**	44.98	**99.99**	**99.99**	**99.61**	**99.54**
	CC	0.20	0.03	0.20	0.21	0.15	0.15
E3	χ^2	0.44	*	0.43	1.33	2.87	**18.49**
	CL	49.1	*	44.71	75.13	91.02	**99.99**
	CC	0.03	*	0.03	0.06	0.09	0.22

* : undefined values preclude the Yates' χ^2-test.
Bold : these succeed the Yates' χ^2-test (at >99% confidence level: $\chi^2 > 6.635$).

tistically significant for only the GT.3 model. The correlations for the other groups have relatively poor confidence levels, ranging from 45% (G3) to 91% (GT.2), so all groups except GT.3 fail to reject the statistical independence hypothesis even at the 95% level ($\chi^2 < 3.841$) – with the possible exception of GT.2 (CL $\approx$ 91%), these other group models are probably not appropriate for predicting the extreme E3 (X) flares. This supports our previous results based on statistical measures alone. We see from Table 4 that the coefficient of contingency, for those groups for which statistical independence has been rejected at the 99% level, ranges from 0.15 to 0.40; these values fall far short of 1.0, which implies that the statistical dependence detected between the observed and forecasted flares is not high. This may be motivation for considering different, perhaps more sophisticated, prediction models.

VI. SUMMARY AND CONCLUSIONS

We compared the statistical results (SR, POD, PC and OP) of the 4 groups of variables (G1–G3, GT) for different data samples depending on flare strength (Events E1–E3). The main results can be summarized as follows.

1. The statistical results of G1 and G3 are nearly the same each other and the 'effect' variables are more reliable predictors than the 'cause' variables.
2. The statistical results of GT are a little worse than those of G1 for weak solar flares, but two results are comparable to each other for strong flares.
3. For the flares weaker than M5 class; all statistical measures show good predictions for all groups. However, stronger flares become difficult to predict well, which is probably due to small statistical samples.
4. The Yates' χ^2 statistical significance tests demonstrate that our statistical results of all flares except for X-class flares were confirmed at the 99% confidence level.
5. Based on our model testing, we recommend a practical strategy for solar X-ray flare predictions. G3 predictions are simpler to make than some other group models, as they are based on just 2 variables (MXF and SF). Thus, we suggest a simple prediction tool based on G3, for each of the 'growing' and 'decaying' phases of sunspot group evolution:

   'Growing' phase:

   $$Y_g \equiv 1.354 + 0.279X_1 + 0.0591X_2, \qquad (8)$$

   'Decaying' phase:

   $$Y_d \equiv 1.345 + 0.302X_1 + 0.0467X_2, \qquad (9)$$

   where $Y_{g,d}$ is the base 10 logarithm of the X-ray flux on the prediction day, X_1 is the base 10 logarithm of the maximum X-ray flux (MXF) on the prior day, and X_2 is the number of strong flares (SF) which took place on the prior day.

Based on these results, we suggest a final recommendation: for quick, reliable forecasts, we provide the simple G3 prediction tool above. The statistical dependence between observed and forecasted flare events is confirmed at greater than the 99% confidence level, except for the extreme E3 (X class) flares. In the case of the less common strong flares, we recommend that GT.1 and GT.3 be used to issue advanced three days warnings of flares stronger than classes M5 and X, respectively. We note that due to the rarity of the strong flares, probabilistic statistical predictions based on mean flaring rate, Poisson statistics and waiting time distributions (see, e.g., Moon *et al.*, 2001; Gallagher, Moon, and Wang, 2002; Wheatland, 2004) may be more appropriate. The past history of occurrence of flares is an important indicator of future flare production (Wheatland, 2004). The autoregressive time-series prediction method might be a convenient tool for a real-time flare forecasting of the advanced three days. It will need further investigation using larger data sets for the strong flares.

ACKNOWLEDGEMENTS

This work is supported by the Korea Research Foundation (ABRL-R14-2002-043-01001-0 and KRF-2005-070-C000059) and the Astrophysical Research Center for the Structure and Evolution of the Cosmos (ARC-SEC) of Korea Science and Engineering Foundation through the Science Research Center (SRC) program.

REFERENCES

Barnes G. and Leka K.D. 2006, 'Photospheric magnetic field properties of flaring versus flare-quiet active regions. III. Magnetic charge topology models, ApJ, 646, 1303–1318

Bartkowiak, A. and Jakimiec, M. 1986, 'Short-term flare activity predictions by means of regression functions calculated for various Zurich class groups observed over the period 1977-1979', P. A. Simon, G. Heckman, M. A. Shea (ed.), *Proceedings of the Workshop on 'Solar-Terrestrial Predictions II'*, Meudon, France, 18–22 June 1984, 285–293

Bartkowiak, A. and Jakimiec, M.: 1990a, 'Short-term predictions of flare activity using alpha-trimmed regression method', Acta Astron. 40, 169–181

Bartkowiak, A. and Jakimiec, M.: 1990b, 'Robust regression with Huber's weights in predictions of flare activity', Acta Astron. 40, 379–388

Bornmann, P. L. and Shaw, D.: 1994, 'Flare rates and the McIntosh active-region classifications', Sol. Phys., 150, 127–146

Bornmann, P. L., Kalmbach, D., and Kulhanek, D.: 1994, 'McIntosh active-region class similarities and suggestions for mergers', Sol. Phys., 150, 147–164

Gallagher, P. T., Moon, Y.-J., and Wang, H.: 2002, 'Active region monitoring and flare forecasting', Sol. Phys., 209, 171–183

Georgoulis, Manolis K. and David M. Rust, 2007, 'Quantitative forecasting of major solar flare', ApJ, 661, 109–112

Hudson, H. S., Haisch, B. M., and Strong, K. T. 1995, 'Comment on "The solar flare myth" by J. T. Gosling', J. Geophys. Res. 100(A3), 3473–3477

Jakimiec, M. and Bartkowiak, A. 1986, 'Relationships among characteristics describing solar active regions', P. A. Simon, G. Heckman, M. A. Shea (ed.), *Proceedings of the Workshop on 'Solar-Terrestrial Predictions II'*, Meudon, France, 18–22 June 1984, 294–299

Jakimiec, M. 1993, 'Two-step regression model used for short-term flare activity prediction', in J. Hruska, M. A. Shea, D. F. Smart, and G. R. Heckman (ed.), *Proceedings of the Workshop on 'Solar-Terrestrial Predictions IV'*, Ottawa, Canada, 22–29 May 1992, v. 2, 180

Jakimiec, M. and Bartkowiak, A. 1994, 'Short-term solar flare predictions by distance-based regression I. Bearalert Regions in 1988 and 1989 – continuous predictors', Acta Astron., 44, 115–140

Kahler, S. W., Sheeley, N. R., Jr., Howard, R. A., Michels, D. J., Koomen, M. J., McGuire, R. E., von Rosenvinge, T. T., and Reames, D. V. 1984, 'Associations between coronal mass ejections and solar energetic proton events', J. Geophys. Res., 89, 9683–9693

Kahler, S. W. 1992, 'Solar flares and coronal mass ejections', ARA&A, 30, 113–141

Langley, R. 1971, *Practical Statistics Simply Explained*, revised ed., Dover Publ., Inc., New York, p. 285

Lee, J. and Kim, K.-S. 1996, 'The prediction of flare production using solar activity data', Publications of Korean Astronomical Society, 11(1), 263–277

Lee, J., Jang, S.-J., Kim, Y.-H., and Kim, K.-S. 1999, 'The prediction of solar activity for solar maximum', Publications of Korean Astronomical Society, 14(2), 103–112

McIntosh, P. S. 1990, 'The classification of sunspot groups', Sol. Phys., 125, 251–267

Moon, Y.-J., Choe, G. S., Yun, H. S., and Park, Y. D. 2001, 'Flaring time interval distribution and spatial correlation of major X-ray solar flares', J. Geophys. Res., 106(A12), 29951–29962

Neidig, D. F., Wiborg, P. H., Seagraves, P. H., Hirman, J. W., and Flowers, W. E. 1986, 'Objective forecasts for solar flares using multivariate discriminant analysis', P. A. Simon, G. Heckman, M. A. Shea (ed.), *Proceedings of the Workshop on 'Solar-Terrestrial Predictions II'*, Meudon, France, 18–22 June 1984, 300–305

Schrijver Carolus J. 2007, 'A characteristic magnetic field pattern associated with all major solar flares and its use in flare forecasting, ApJ, 655, 117–120

Wheatland M.S. 2000, 'The origin of the solar flare waiting-time distribution', ApJ, 536, 109–112

Wheatland M.S. 2004, 'A bayesian approach to solar flare prediction', ApJ, 609, 1134–1139

Wilson, R. M. 1994, 'A comment on the suspected solar neutrino-solar activity connection', Sol. Phys., 149, 391–394