Predicting Solar Energetic Proton Events (E > 10 MeV)

Marlon Núñez
Universidad de Málaga

Abstract: A high level of proton radiation exposure can be dangerous to astronauts, satellite equipment and air passengers/crew flying along polar routes. The presented solar energetic proton (SEP) event forecaster is based on a dual-model approach for predicting the time interval within which the integral proton flux is expected to meet or surpass the Space Weather Prediction Center threshold of J (E >10 MeV) = 10 pr cm^{-2} sr^{-1} s^{-1} and the intensity of the first hours of well- and poorly-connected SEP events. This forecaster analyzes flare and near-Earth space environment data (soft X-ray, differential and integral proton fluxes). The purpose of the first model is to identify precursors of well-connected events by empirically estimating the magnetic connectivity from the associated CME/flare process zone to the near-Earth environment and identifying the flare temporally associated with the phenomenon. The goal of the second model is to identify precursors of poorly-connected events by using a regression model that checks whether the differential proton flux behavior is similar to that in the beginning phases of previous historically poorly-connected SEP events, and thus deduce similar consequences. An additional module applies a higher-level analysis for inferring additional information about the situation, by filtering out inconsistent preliminary forecasts and estimating the intensity of the first hours of the predicted SEP events. The high-level module periodically retrieves solar data and, in the case of well-connected events, automatically identifies the associated flare and active region. For the events of solar cycles 22 and 23 of the NOAA/SWPC SEP list, the presented dual-model system, called UMASEP, has a probability of detection of all well- and poorly-connected events of 80.72% (134/166) and a false alarm rate of 33.99% (69/203), which outperforms current automatic forecasters in predicting >10 MeV SEP events. The presented forecaster has an average warning time of 5 h 10 min for the successfully predicted events, 1h 5 min for well-connected events and 8 h 28 min for poorly-connected events, with a maximum warning time of 24 hours for very gradual SEP events.

1. Introduction

A solar radiation storm occurs when a major event on the Sun or in the interplanetary medium accelerates protons toward Earth. These events have a significant effect on satellites and humans [*Lucci et al.,* 2005]. In the atmosphere, they interact to produce penetrating neutrons that irradiate passengers and flight crews in commercial aircraft flying polar routes [*Beck et al.,* 2005]. In space, protons affect electronic circuits, solar cells and the mirrored surfaces of satellites and spacecraft [*Koskinen et al.,* 2001], and they may penetrate the space suits of astronauts aboard the International Space Station [*Miller et al.,* 2003]. A high level of proton radiation exposure can be experienced by astronauts who are outside a protective shield in space [*Lucci et al.,* 2005]. In the future, solar proton events will be important factors to take into account in interplanetary flights to the Moon, the asteroid belt, and Mars [*Rapp,* 2006].

Space weather systems need to predict radiation events early and reliably. In particular, they should neither miss radiation events nor issue false warnings at an unacceptably high rate that might be disruptive for space activities [*Kahler et al.,* 2007]. SEP forecasts should enable users to take precautions immediately upon advance warning messages until the radiation event has ceased. The more anticipation there is, the lower are the risk to health and risk of damage to equipment.

Traditionally, SEPs have been difficult to forecast. In 2004, Donald Reames of the Goddard Space Center wrote [*Reames,* 2004]: "Our ability to predict SEP events is almost nonexistent, but that does not prevent predictions from being made. …reliable predictions of the onset and fluence of an SEP event prior to its occurrence are not likely in our lifetime". According to the evaluation of historical data of solar cycles 22 and 23 (from September 1986 to December 2007), the presented system, called UMASEP, has a probability of detection of SEPs (E > 10 MeV) of 80.72% (134/166), a false alarm rate of 33.99% (69/203), and an average warning time of 5 h 10 min, which outperforms current automatic forecasters and allows spacecraft and aircraft operators to take better preventive actions against hazardous proton radiation events. Warning time is the temporal distance between the time when the prediction is issued and the time when the integral proton flux meets or surpasses the Space Weather Prediction Center (SWPC) SEP threshold of J (E >10 MeV) = 10 proton flux units (1 pfu = 1 pr cm^{-2} sr^{-1} s^{-1}).

To make its predictions, UMASEP analyzes soft X-ray flux, differential proton flux (E= 9-500 MeV) and integral proton flux (E>10 MeV) data. This system does not analyze data for estimating CME-driven shock properties (e.g., by using type II radio emissions). UMASEP is compared with other well-known and operative forecasters [*Balch,* 2008; *Kahler et al.,* 2007; *Laurenza et al.,* 2009; *Posner,* 2007], including a preliminary version of the presented system [*Núñez,* 2009; *Núñez and Núñez,* 2009]. UMASEP's forecasts are currently downloaded every 5 minutes from users' space weather systems, such as the iSWA system [*Maddox et al.,* 2008] of NASA (http://iswa.gsfc.nasa.gov), the SEISOP system [*Di Marco et al,.* 2008] of the European Space Agency (ESA), and the European Space Weather Portal http://www.spaceweather.eu/en/forecast/uma_sep.

Section 2 introduces the system architecture and presents the forecasting models. Section 3 describes the outputs of UMASEP and presents its statistical performance on historical activity data from solar cycles 22 and 23. Section 4 shows the statistical performance of this approach. Section 5 compares the presented forecaster with other current forecasters. Finally, Section 6 presents some future directions and conclusions.

2. Solar protons, magnetic field lines, and forecasting models

Embedded in the solar wind plasma is the heliospheric magnetic field that becomes the interplanetary magnetic field (IMF). The radial outflow of the solar wind from the corona transports the IMF into interplanetary space, while the footprint of the field line remains anchored in the solar atmosphere. The combination of the solar wind outflow and the Sun's rotation results in the magnetic fields having an Archimedean spiral configuration [*Parker,* 1958]. Solar protons are thus thought to be guided by spiral field lines in an average solar wind. Depending on the observer's longitude relative to the originating solar event, several intensity profiles are possible [*Reames,* 2004]. These profiles depend on the magnetic connectivity between the observer (e.g., satellite) and the solar parent event:

- If the satellite, situated near Earth (at 1 AU), is magnetically well connected with the originating solar event, it may observe rapidly rising solar proton intensities (E>10 MeV). The particles in these SEP events are accelerated by flares and coronal mass ejections (CMEs). Mostly associated with western (right-hand side of the Sun) events, well-connected SEP events may also be associated with solar eastern events.

- If the satellite, situated at 1 AU, is magnetically poorly connected with the originating solar event, it may observe a slow increment in proton intensity (E>10 MeV), surpassing the SEP threshold[1] 1-3 days after the solar event. The particles in these SEP events are accelerated by shock waves driven out from the Sun by CMEs. The observer witnesses a maximum intensity only after crossing through the shock into the region where field lines connect to the shock nose from behind.

There are numerous situations in which the event is neither purely "well" nor "poorly" - connected. For example, there are cases of poorly connected situations during which new solar events inject a new population of energetic particles into the region in and behind a CME [*Reames,* 2004; *Cane et al.,* 2003].

The prediction of SEP events (E>10 MeV) can be conceived as a synergy of models that forecast different types of SEP events. UMASEP, presented in this paper, is based on two models and an additional high-level module, as presented in Figure 1. This figure also illustrates the inputs and outputs of the aforementioned components.

The Well-Connected SEP Forecasting model (see Figure 1) tries to identify precursors of well-connected events by empirically estimating the magnetic connectivity from the associated CME/flare process zone to the near-Earth environment and identifying a great flare temporally associated with the phenomenon, independently of the flare's heliolongitude. This model analyzes flare and near-Earth space environment data (soft X-ray and differential proton fluxes in the range 9-500 MeV) and fires a warning when it determines that a magnetic connection is present and that the associated

[1] Proton fluxes are integral 5-minute averages for energies >10 MeV, given in Particle Flux Units (pfu), measured by a spacecraft (i.e. GOES) at Geosynchronous orbit: 1 pfu = 1 p cm-2 sr-1 s-1. The start of an SEP event is the first of 3 consecutive data points with fluxes greater than or equal to 10 pfu, which is considered the SEP threshold.

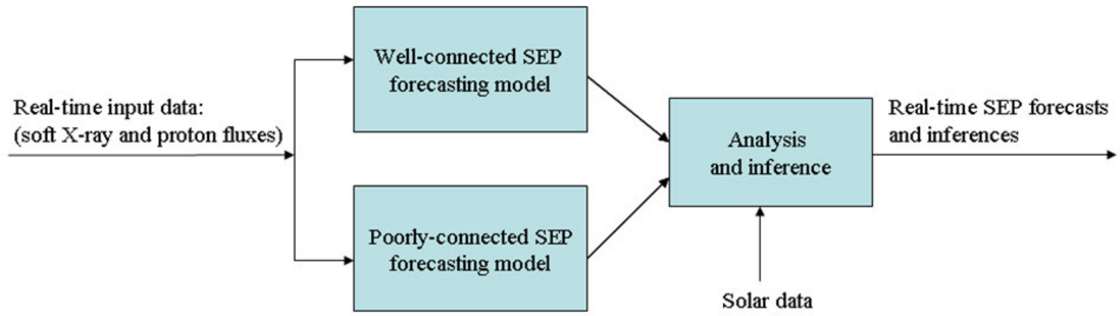

Figure 1. Main modules of the UMASEP system for predicting SEP events (E>10 MeV).

flare is greater than 7×10^{-6} W m^{-2} (C7 flares[2]). Section 2.1 describes this model in detail.

The Poorly-Connected SEP Forecasting model (see Figure 1) tries to identify precursors of poorly-connected events, using a regression model that checks whether the differential proton flux behavior is similar to the *beginning* phase of previous historically poorly-connected SEP events and thus deduces similar consequences. In a pure poorly connected event, the beginning phase corresponds to the first small proton flux enhancements, during the interval of time in which we assume that the observer is beginning to be connected to the shock; that is, there are proton enhancements, but the integral proton flux has not met or surpassed the SWPC SEP threshold. This model does not predict the peak SEP time, that is, the time when the observer is well connected to the shock nose from behind. If the model recognizes that a poorly-connected event is in progress, it sends a preliminary forecast. The regression model was trained offline with proton flux data on previous historically poorly-connected SEP events from solar cycles 22 and 23. Section 2.2 describes this model in detail.

The Analysis and Inference module receives preliminary forecasts from the aforementioned models, as shown in Figure 1, and performs a higher-level analysis to infer additional information about the situation and filter out non-consistent preliminary forecasts. This module also estimates the intensity of the first hours of the proton event. Only in the case of well-connected event forecasts, this high-level module periodically retrieves solar data (if available) and automatically identifies the associated flare and active region. Section 2.3 describes this module in detail.

[2] Solar flares are classified according to their X-ray brightness in the wavelength range 1 to 8 Angstroms. There are X-class flares, which are big ($I >= 10^{-4}$ Watts m^{-2}), M-class flares are medium-sized ($10^{-5} <= I < 10^{-4}$ Watts m^{-2}); C-class flares are small $10^{-6} <= I < 10^{-5}$ Watts m^{-2}). Each category of X-ray flares has nine subdivisions ranging from, e.g., C1 to C9, M1 to M9, and X1 to X9.

2.1. Description of the well-connected SEP forecasting model

Magnetically well-connected SEP events may be detected at 1 AU several minutes or hours after the originating solar event. The energetic solar particles follow the spiral paths all the way. If the Parker spiral were stable, we would know the magnetic connection from the observer to the associated CME/flare zone situated at approximately W60; however, solar particles depend on non-stable magnetic field lines “frozen” in a solar wind that is highly variable in both time and space.

Current operational models are empirical and perform real-time analysis of signals associated with the phenomena. Section 2.1.1 includes a brief summary of each of these empirical systems, and Section 5 provides a comparison with the presented forecaster.

There is not a purely physics-based model for predicting well-connected SEP events in real-time. The ISWA system [*Maddox et al.*, 2008] presents an approach for estimating the Sun-Earth connection that uses a combination of physics-based and empirical models of the corona and heliosphere (CorHel [*Linker et al*., 2010], WSA [*Arge and Pizzo*, 2000] and ENLIL [*Xie et al.*, 2004]) to calculate positions of magnetic fields connected to the Earth. Currently, the ISWA’s approach does not explicitly predict well-connected events but, it could be used as a tool to do so in the future.

SEP events are highly variable in their spectral characteristics and elemental composition, as studied by *Tylka et al.* [2005]. They proposed that this variability arises from the interplay of two factors: a compound seed population, typically comprising both solar wind and flare suprathermals, and the shock geometry. Quasi-perpendicular shocks preferentially reaccelerate seed particles from flares. Also, near well-connected longitudes, the strongest parts of the shock are more likely to intercept the Sun-Earth field line while still near the Sun. Physics-based numerical models, mainly based on SEP acceleration at shocks, have been developed for simulating and predicting well-connected SEP events (as well as poorly-connected events) [e.g., *Sokolov et al.,* 2004; *Aran et al.,* 2006], however they are not yet operational.

There is no consensus regarding the role of the associated flare in the acceleration process of >10 MeV protons. *Tylka et al.* [2005] mentioned that in well-connected events, it is possible, although not yet proven, that the associated flare may contribute seed particles if open field lines connect the flare site to the shock. Therefore, these ‘‘fresh’’ (as opposed to ‘‘remnant’’) flare suprathermals might also explain the comparatively high proportion of Fe-rich events at well-connected longitudes. *Cane et al.* [2006] suggested that variations in the elemental composition of SEP events mainly arise from the combination of flare particles and shock acceleration of these particles and/or the ambient medium. Other authors [*Klein et al*., 2005; *Marque et al.,* 2006] have suggested a more important role for flares in the CME/flare scenario. Although there is no consensus on this topic, there is no doubt that the associated flare is important in well-connected SEP events.

2.1.1. Magnetic connection estimation by correlating soft X-rays and differential proton fluxes

All empirical and operational SEP forecasting methods, including the well-connected SEP forecasting model presented in this section, currently rely more on data about the associated flare rather than the associated CME-driven shock to predict well-connected SEP events [*Laurenza et al.,* 2009; *Balch*, 2008; *Posner*, 2007; *Kahler et al.,* 2007]. The method proposed by *Laurenza et al*. [2009] for predicting well-connected SEP events is based on flare location, flare size, and evidence of particle acceleration/escape as parameterized by flare longitude, time integrated soft X-ray intensity, and time-integrated intensity of type III radio emission at 1 MHz, respectively. The method proposed by *Balch* [2008] assumes that there is a relationship between the intensity of solar flare emissions and SEP event occurrence. Balch’s method is based on the soft X-ray peak flux and time-integrated flux, the occurrence or non-occurrence of type II (associated with CME-driven shocks) and/or type IV radio bursts, and the Hα flare location. *Kahler et al.* [2007] developed a method for predicting SEP events by analyzing the solar flare peak, time-integrated X-ray, radio fluxes, times of onset and maxima, and solar flare locations. *Posner* [2007] developed an electron-based SEP prediction technique that exploits the shorter transit time of electrons (mainly accelerated by the associated flare) relative to ions. *Belov* [2009] proposed a method that calculates the probability of SEP events from X-ray observations.

Most current operational SEP forecasters, including our system, also assume that the associated flare is important for predicting well-connected SEP events. Our system, like most current systems (with the exception of Balch’s), has the limitation of not analyzing acceleration signatures at the shock waves (e.g., type II radio emissions). The analysis of flare-associated data seems to be simpler and effective in the complex task of predicting SEP events.

We determine whether particles have escaped along IMF field lines to the observer by correlating X-rays with proton flux data. Our well-connected SEP forecasting model assumes that the observer site is connected with the heliolongitude of the associated CME/flare process zone when the X-ray flux (from the associated flare) is somehow correlated with at least one of the differential proton channels at the observer site. Our assumption is that a magnetic connection is occurring when there is a high and lasting correlation between the sequence (or time series) of first derivatives of the X-ray flux and the first derivatives of at least one differential proton flux, which was studied in the first version of the presented forecaster [*Núñez et al*., 2006]. By correlating these signals, we may deduce that the accelerated particles that reach the observer are somehow associated with the X-ray process. To show this correlation with real data, the chart at the top of Figure 2a shows the first derivatives of soft X-rays and a differential proton channel corresponding to the 80-165 MeV energies (or P6 in the NOAA terminology) on October 26, 2003. The first derivatives were normalized to the maximum value of each time series for the shown time interval. At 18:00 UT, a very similar evolution of the aforementioned time series is observed, which is probably due to a magnetic connection from the CME/flare heliolongitude to the observer. Note that the strongest P6 flux derivatives occurred 30 minutes after the strongest X-ray flux derivatives.

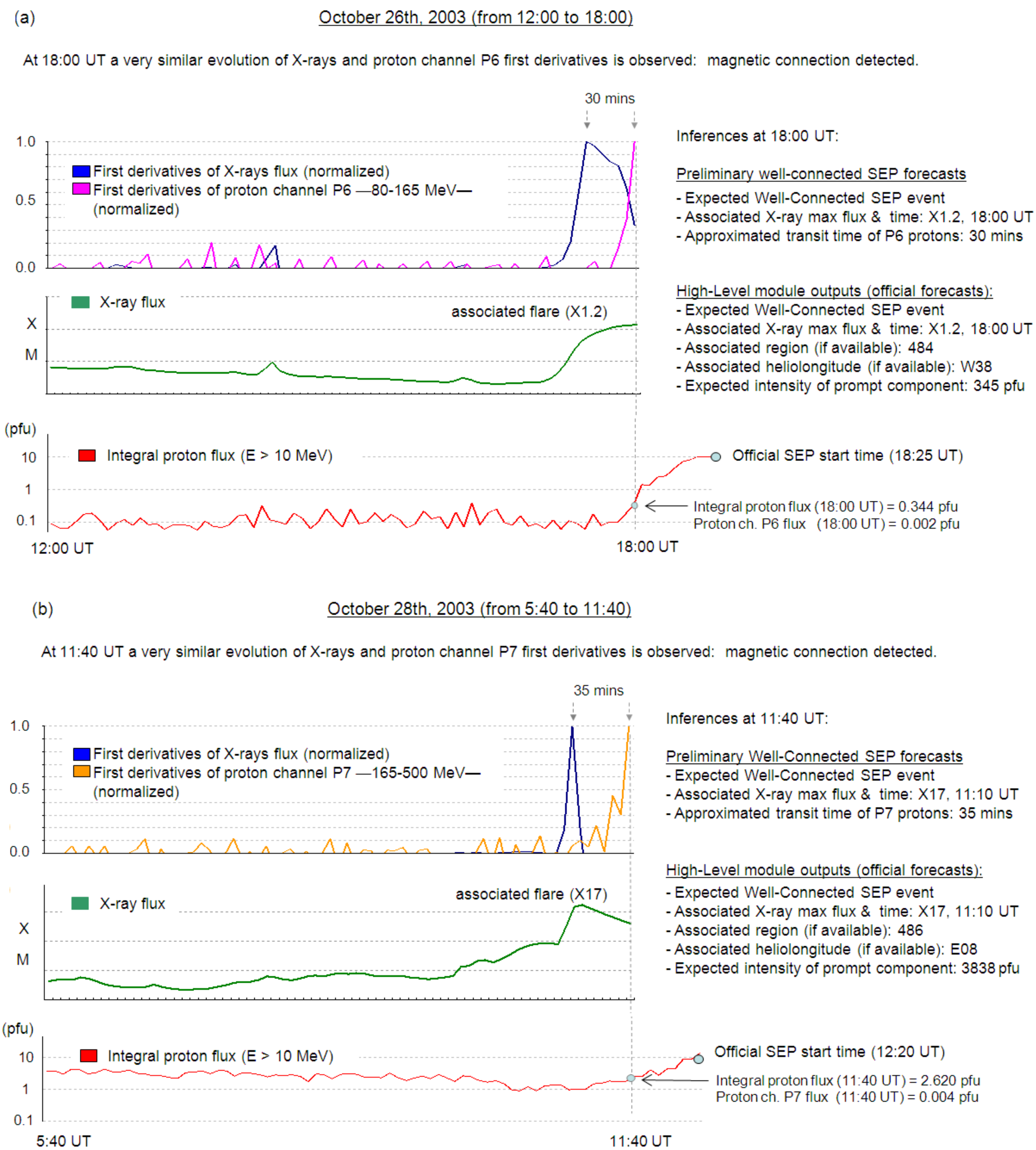

Figure 2. This figure illustrates the functioning of UMASEP for predicting well-connected SEP events. Figure 2a shows the evolution of the normalized first derivatives of soft X-rays and a differential proton channel immediately before 18:00 on October 26, 2003. The normalization of the derivatives was carried out with regard to the maximum value of the shown time interval. Because the evolutions of both time series are very similar, the system infers that this pattern is due to a magnetic connection from the solar parent event location to the observer (GOES-10). After the analysis of the situation, the system predicted a well-connected SEP event (see Figure 4) and found that it was associated with a western flare; however, its heliolongitude was not taken into account during the forecasting process. Figure 2b shows the same analysis for predicting the SEP event of October 28, 2003, which was associated with an eastern flare.

Figure 2b shows the same correlation for another well-connected SEP event that occurred on October 28, 2003. The rise in the first derivatives of the P7 channel occurred 35 minutes after the strongest first derivatives of the X-ray process.

Figure 2a also shows that UMASEP may issue a well-connected SEP forecast when no enhancement in the integral proton flux (E> 10 MeV) is observed. In the case of October 26, 2003, the forecaster issued the prediction when the integral proton flux was 0.344 pfu, which is much lower than the SEP threshold of 10 pfu (see lower chart of Figure 2a; see also Figure 4). The low levels of the integral proton flux at the time of the prediction arose because UMASEP analyzes differential proton channels, regardless of the magnitude of their fluxes. It may find a lasting rise in the first derivatives of a differential proton channel (even in the level of very low fluxes), correlate it with a lasting rise in the first derivatives of the soft X-ray flux (even at low levels), and identify a magnetic connection. For example, the lower right area of Figure 2a shows that the flux of the correlated channel (P6) was very low (0.002 pfu according to the GOES-10 satellite) at the time of the prediction. In the case of October 28, 2003, the forecaster issued the prediction when the integral proton flux was 2.62 pfu (see lower chart of Figure 2b), which is near 10 pfu; however, the integral flux was at those levels (ranging from 1 to 4 pfu) during the previous six hours, and therefore 2.62 pfu was not the highest flux during the observed period.

Prior to most of the prompt SEP events of solar cycles 22 and 23, the same correlation occurs, as shown in Figure 2. To measure the correlation between both time series, several methods might be applied. We experimented with classic lag-correlation functions [*Wei*, 2005], which measure the strength of the relationship between two time series in which an unknown lag between both time series is present; however, we decided to design an ad-hoc correlation measure that yielded better results for our purposes. This measure is presented in the next section. If there is no correlation between the aforementioned time series, the model concludes that there is no magnetic connection and does not issue any well-connected SEP forecast, regardless of the intensity or heliolongitude of the observed flares. During December 13 and 14, 2001, for instance, there were six M flares and one X flare. Because the system did not detect any magnetic connection, it successfully predicted that no event would occur. That is, the magnetic connection is a necessary condition for predicting well-connected events.

We calculate a bi-series correlation between the normalized first derivatives of the soft X-ray (SXR) flux of the primary GOES and the normalized first derivatives of the flux of each differential proton channel (9 MeV<E< 500 MeV) of every available GOES. The magnetic connectivity estimation will be the greatest correlation value found after processing all available GOES satellite data. If the connectivity is lasting and there is a >C7 flare temporally associated with the phenomenon, the well-connected SEP forecasting model predicts an event. Although they are a minority, well-connected eastern SEP events are possible (e.g., October 28, 2003, in Figure 2b). For this reason, our approach does not filter out forecasts according to a flare's heliolongitude.

2.1.2. A correlation approach using an ad-hoc similarity measure

To identify a possible magnetic connection between a CME/flare process location and the Earth, a set of lag correlations has to be measured between the soft X-ray flux and each of the five differential proton fluxes (9 MeV < E < 500 MeV) measured at 1 AU by all the available GOES satellites. At every time step t, the model performs five lag correlations for each satellite and selects the highest correlation as the basis of the conclusions.

The well-connected SEP forecasting model transforms the original X-ray time series, say $A=\{a_0, a_1, .., a_n\}$, and the differential proton channel $B=\{b_0, b_1, ..b_n\}$ to a time series of normalized first derivatives on which the correlation analysis is applied. We construct DA and DB by calculating their derivatives. That is, $DA = \{da_1, da_2, .., da_n\}$ and $DB= \{db_1, db_2, .., db_n\}$, where $da_i = a_i - a_{i-1}$ and $db_i = b_i - b_{i-1}$. At every time t, two subsequences sDA and sDB of length l are constructed with the derivatives of DA and DB normalized to $maxsDA$ and $maxsDB$, which are the maximum values of sDA and sDB in the interval from t-l to t, respectively. The transformed time series are $sDA= \{da_{t-l}/maxsDA, da_{t-l+1}/maxsDA, .., da_t/maxsDA\}$ and $sDB= \{db_{t-l}/maxsDB, db_{t-l+1}/maxsDB, .., db_t/maxsDB\}$.

Then, the approach may identify potential cause-consequence pairs in sDA and sDB, with the possible causes being the X-ray flux first derivatives, also called X-ray *fluctuations*, and the possible consequences the differential proton flux *fluctuations*, also called proton fluctuations. A flux fluctuation is analyzed only if it surpasses a normalized threshold $h \in [0, 1]$, where 1 is the highest flux fluctuation (see Figure 2) in the interval from *t-l* to *t*. A pair is discarded if the time between the causative fluctuation and the consequential fluctuation is greater than eight and a half minutes, which is the time that a solar particle takes to reach the Earth traveling at the speed of light. Because there are several ways to pair X-ray fluctuations to differential proton fluctuations, the approach collects all possible combinations of consecutive cause-consequence pairs. Every combination is called a *CCsequence.*

To calculate the fluctuation similarity, each potential *CCsequence* has a set of possible cause-consequence pairs. For every pair $p = \{i, j\}$, where i is an X-ray *fluctuation* and j is a differential proton *fluctuation*, there is a $temporalDifference_p$ that is, $|time(i) - time(j)|$, and an $intensityDifference_p$, that is, $|intensity(i)\text{-}intensity(j)|$. A cause-effect pattern between *sDA* and *sDB* is found when there is a sequence of pairs with very similar *temporalDifferences* and *intensityDifferences*. To measure the similarity between the analyzed subsequences, we used an ad-hoc formula:

$$fluctuationSimilarity(CCsequence) = w_{tempD}\frac{\mu_{temD}+\varepsilon}{\mu_{temD}+\sigma_{temD}+\varepsilon} + w_{\mathrm{int}\,D}\frac{\mu_{intD}+\varepsilon}{\mu_{\mathrm{int}\,D}+\sigma_{\mathrm{int}\,D}+\varepsilon} \quad (1)$$

where w_{temD} and w_{intD} are weights of the similarity in terms of temporal and intensity differences, respectively; μ_{temD} and σ_{temD} are the average and the standard deviation of the *temporalDifferences* of the pairs within *CCsequence*; μ_{intD} and σ_{intD} are the average and the standard deviation of the *temporalDifferences* and *intensityDifference*s of the pairs within *CCsequence*; and ε is a very low value used to avoid possible divisions by 0.

The well-connected model calculates the *fluctuationSimilarity* for all differential proton channels and identifies the highest fluctuation similarity *fs*. Finally, the method can issue the following conclusions:

- If the *fluctuation similarity fs* is lower than a fluctuation-similarity threshold *m*, there is no magnetic connection associated with any of the analyzed proton channels, and therefore no well-connected SEP forecast is issued.

- If the *fluctuation similarity fs* is greater than or equal to the fluctuation-similarity threshold *m*, two conclusions are issued: there is a magnetic connection whose (normalized) strength is *fs*, and the average of the temporal distances between the causes and consequences within *CCsequence* is the estimated transit time of protons from the Sun to 1 AU. The associated flare may be identified in the information within *CCsequence*. The highest original (X-ray) flux of the corresponding causative fluctuations in pairs within *CCsequence* corresponds to the peak of the associated flare. If the peak of the associated flare is greater than a certain X-ray flux threshold *f,* then a preliminary well-connected SEP forecast is sent to the Analysis and Inference Module, including the time and X-ray peak of the associated flare.

Because we had to increment the probability of detection (POD) and reduce the false alarm rate (FAR)[3], we searched for an optimal configuration of the weights w_{temD} and w_{intD} (factors of the similarity function), the parameter *l* (length of the analyzed time interval), and thresholds *h* (minimum height of the normalized fluctuations), *f* (minimum value of the X-ray flux of the associated flare to issue a forecast), and *m* (minimum fluctuation similarity) while predicting well-connected SEP forecasts during solar cycles 22 and 23. The threshold *f* is the minimum flare intensity for issuing a preliminary forecast, which is explained in Section 2.1.3. A *general forecasting performance* measure was needed to find the optimal configuration. We used $w_{precision}\cdot Precision+w_{recall}\cdot Recall$, where *Recall* is the POD and *Precision* is 1 – FAR (*Davis and Goadrich* [2006]), and $w_{\mathrm{precision}}$ and w_{recall} are weights. With these types of multi-objective problems, designers usually give more weight to one objective than to the other. We decided to give equal importance to POD and 1-FAR; therefore, the weights are 0.5.

To find a highly effective configuration (not necessarily the best one) of weights, parameters and thresholds, we used a multi-resolution optimization. That is, we first searched the two most optimal threshold configurations using low-resolution steps. For every configuration found, we applied a new search by using higher resolution steps in the neighborhoods of the solutions found in the previous phase. The width of the new range for every threshold/weight (to be optimized using higher-resolution steps) was a tenth of the original low-

[3] The forecaster performance is evaluated in terms of Probability of Detection (POD) as A/(A + C) and False Alarm Rate (FAR) as = B/(A + B), where A is the number of correct forecasts (an SEP event was forecast and one occurred), B is the number of false alarms (an SEP event was forecast but none occurred) and C is the number of missed events (no SEP event was predicted but an event did occur).

resolution width. We repeated the process until a stable *general forecasting performance* was reached.

2.1.3. Issuing preliminary forecasts of well-connected SEP events

The existence of a magnetic connection is a necessary condition to forecast a well-connected SEP; it shows that particles are escaping from the CME/flare process and are arriving at 1 AU along magnetic field lines. In addition to checking the existence of a magnetic connection, the model also needs to check whether the flux peak of the associated flare surpasses the threshold *f*. The precision-recall method mentioned at the end of Section 2.1.2 shows that *f* is $7x10^{-6}$ W m^{-2}, which is C7 in terms of the soft X-ray flux. In summary, the model concludes that the integral proton flux will meet or surpass 10 pfu (official SWPC SEP threshold) due to a well-connected event when a magnetic connection is detected (according to the procedure in Section 2.1.2) and the encountered associated flare has a soft X-ray flux of C7 or greater. Note that this forecasting model does not need to check the heliolongitude of the associated flare.

A preliminary well-connected SEP forecast is sent to the Analysis and Inference Module, including the time and X-ray peak of the associated flare. This module analyzes the information regarding the preliminary forecast. If a preliminary forecast is not filtered out by the Analysis and Inference Module, the module calculates the expected intensity of the predicted SEP and sends all this information and other inferences to the user.

The satisfactory results of the well-connected SEP forecasting model (see Section 4) and its necessary condition of the existence of magnetic connections between the observer and the X-ray process might imply that there also exist intermediate magnetic connections from the flare site to the CME-driven shock (where protons are finally reaccelerated) during the first hours of the proton enhancement. Our results might also be explained by taking into account other solar processes (e.g., Chen and Kunkel [2010] found evidence that there is a physical relationship between flare energy release and poloidal magnetic flux injection in CMEs).

2.2. Description of the poorly-connected SEP forecasting model

If a satellite situated at 1 AU is magnetically poorly connected to the originating solar event, it may detect a slow increment in proton intensity (E>10 MeV), surpassing the SEP threshold 1-3 days after the solar event. These types of SEPs have been widely characterized and simulated by a combination of solar corona, solar wind and particle transport models [*Lionello et al.,* 2003; *Odstrcil et al.,* 2004; *Toth et al., 2005*; *Schwadron et al.,* 2010; *Aran et al*., 2006]. Most of the models are based on MHD equations, which depend on the initial conditions. Nowadays, the uncertainty in the initial conditions of each individual model may be the main reason why the combined models are not yet operational; however, they offer the most logical approach for the real-time prediction of all types of SEPs, particularly poorly connected events, in the future.

Four known empirical SEP forecasters that are able to predict poorly connected events are PROTONS [*Balch,* 2008], PPS [*Kahler et al.,* 2007], and the models developed by *Laurenza et al.*[2009] and Posner [2007]. These systems were summarized in Section 2.1.1. Section 5.1 shows the results of the comparison of these two systems with UMASEP.

The method presented in Section 2.1 for well-connected SEP forecasting is not adequate for predicting poorly-connected SEP events, because in these events, there is no magnetic connection from the parent solar event's heliolongitude to the observer. In poorly-connected events, the particles are accelerated by interplanetary CME-driven shocks or CIR (corotating interaction region) events whose locations and evolutions are difficult to estimate. Because there is considerable uncertainty regarding the source of particle acceleration, we propose a different approach.

In the case of poorly-connected SEP forecasting, we assume that a lasting gradual rise in several differential proton fluxes, not necessarily all —that is, not necessarily the integral proton flux—, is a symptom that the observer is beginning to be connected to the shock. We had to analyze five GOES differential proton channels, P3 to P7 (see note b in Table 1 for more detailed energy range information). It is not clear when a joint rise ends up with an integral proton flux surpassing 10 pfu, or when it ends up with an integral proton flux decreasing. On the other hand, there are multiple possibilities of joint rises: (P3, P4, P5), (P4, P5, P6), (P5, P6 P7), (P3, P4, P5, P6), (P3, P4), etc. We do not know which sets of differential proton channels are the more probable, or if they have to be consecutive, or which behavior would give symptoms of continuity or symptoms of joint decrement. When the analysis is so hard, and no relationships are clear data mining is a good

Table 1. Fields of the Training Table Needed to Construct a Model Tree

Fields[a]	Proton Flux Time Series
$\log P3(t)$[b]	$\log P3, \ldots, \log P7$[c]
$\log P3(t\text{-}1)$[b]	$\log P3, \ldots, \log P7$[c]
… $\log P3(t\text{-}12)$[b]	$\log P3, \ldots, \log P7$[c]
…[b]	$\log P3, \ldots, \log P7$[c]
$\log P7(t)$[b]	$\log P3, \ldots, \log P7$[c]
$\log P7(t\text{-}1)$[b]	$\log P3, \ldots, \log P7$[c]
… $\log P7(t\text{-}12)$[b]	$\log P3, \ldots, \log P7$[c]
$D\log P3(t)$[b]	$D\log P3, \ldots, D\log P7$[d]
$D\log P3(t\text{-}1)$[b]	$D\log P3, \ldots, D\log P7$[d]
… $\log DP3(t\text{-}12)$[b]	$D\log P3, \ldots, D\log P7$[d]
$D\log P7(t)$[b]	$D\log P3, \ldots, D\log P7$[d]
$D\log P7(t\text{-}1)$[b]	$D\log P3, \ldots, D\log P7$[d]
… $D\log P7(t\text{-}12)$[b]	$D\log P3, \ldots, D\log P7$[d]
$D\log P3(t) + D\log P3(t\text{-}1))/2$[e]	$\log P3$, $\log P4$, …, $\log P7$, $D\log P3$, …, $D\log P7$, $\log Ipf$, and $D\log Ipf$[f]
$D\log P4(t) + D\log P4(t\text{-}1))/2$[e]	$\log P3$, $\log P4$, …, $\log P7$, $D\log P3$, …, $D\log P7$, $\log Ipf$, and $D\log Ipf$[f]
$D\log P7(t) + D\log P7(t\text{-}1) + D\log P7(t\text{-}2))/3$[e]	$\log P3$, $\log P4$, …, $\log P7$, $D\log P3$, …, $D\log P7$, $\log Ipf$, and $D\log Ipf$[f]
…[e]	$\log P3$, $\log P4$, …, $\log P7$, $D\log P3$, …, $D\log P7$, $\log Ipf$, and $D\log Ipf$[f]
$\log Ipf(t)$[b]	$\log Ipf$ and $D\log Ipf$[g]
$\log Ipf(t\text{-}1)$[b]	$\log Ipf$ and $D\log Ipf$[g]
… $\log Ipf(t\text{-}12)$[b]	$\log Ipf$ and $D\log Ipf$[g]
$D\log Ipf(t)$[b]	$\log Ipf$ and $D\log Ipf$[g]
$D\log Ipf(t\text{-}1)$[b]	$\log Ipf$ and $D\log Ipf$[g]
… $D\log Ipf(t\text{-}12)$[b]	$\log Ipf$ and $D\log Ipf$[g]
$\log Ipf(t + i)$[h]	$\log Ipf$[i]

[a]Past and future values on an hourly basis.
[b]There are 160 input temporal variables (related to the past).
[c]Proton flux time series, where P3…P7 are differential proton channels: $P3$ is the time series of flux J ($9 < E < 15$ MeV) pr cm^{-2} sr^{-1} s^{-1}, $P4$ is the time series of flux J $15 < E < 40$ MeV) pr cm^{-2} sr^{-1} s^{-1}, $P5$ is the time series of flux J ($40 < E < 80$ MeV) pr cm^{-2} sr^{-1} s^{-1}, $P6$ is the time series of flux J ($80 < E < 165$ MeV) pr cm^{-2} sr^{-1} s^{-1}, $P7$ is the time series of flux J ($165 < E < 500$ MeV) pr cm^{-2} sr^{-1} s^{-1}, Ipf is the time series that represents the integral proton flux J ($E > 10$ MeV) pr cm^{-2} sr^{-1} s^{-1}, $\log X$ is the time series based on the base-10 logarithm of the flux X (i.e., $\log P3$). Twelve past values each of the proton channel fluxes.
[d]Proton flux time series. DX is the time series based on the first derivatives of the flux X (i.e., $D\log P3$). Twelve past values of the first derivatives of each of the proton channel fluxes.
[e]There are 160 input temporal variables (related to the past). Sixteen calculated variables.
[f]Sixteen calculated variables.
[g]Proton flux time series. Twelve past values of the integral proton channel fluxes and their derivatives.
[h]Response variable (related to the future).
[i]Predicted future value at $t + i$ hours of the integral proton flux, where $i = 1..24$.

solution. Therefore we decided to train a regression system with past data to predict when a joint rise of differential proton channels would end up becoming a poorly-connected SEP and when those joint rises would end up decreasing.

Data mining has methods that allow us to find clear relationships while predicting numerical classes (e.g., regression trees [*Breiman et al.*, 1984]), however they are not the most accurate. There are data mining methods that are very accurate; unfortunately they are obscure, called "black box" methods (e.g., neural nets, ensembles). Since our approach will be evaluated by its accuracy, not by the understandability of the found relationships, we decided to use an ensemble of regression models, that is, a regression ensemble, composed of several sub-models, with past data on solar cycles 22 and 23.

Data mining modeling is based on the advances of computational intelligence, and has proven to be a powerful approach to a number of problems in several domains, including space weather [*Qahwaji and Colak,* 2007; *Boberg et al.,* 2000; *Nuñez et al.,* 2005]. In the field of data mining, several learning algorithms have been proposed to construct single models; however, it has been widely shown that a combination of models, which is called ensemble modeling [*Hansen and Salamon,* 1990], yields better results than using individual models.

In summary, we designed a purely empirical approach focused on analyzing whether the differential proton flux behavior is similar to that of the *beginning phase* of previous historically poorly-connected SEP events, and from that information, deducing similar consequences. In a pure poorly-connected event, the beginning phase corresponds to the interval time in which we assume that the observer is beginning to be connected to the shock; that is, there are small proton enhancements. The goal of the poorly-connected model is to predict the time interval within which the integral proton flux is expected to meet or surpass the SWPC SEP threshold and the intensity of the first hours of the SEP event. This model does not predict the peak SEP time, that is, the time when the observer is fully (or well-) connected to the shock nose from behind.

2.2.1. Nonlinear regression and ensemble of regression models

Linear regression models help to explain observations of a dependent variable, usually denoted by y, using observed values of m independent variables, usually denoted by x_1, x_2, ..., x_m. That is, $y = \varepsilon + k_1x_1 + k_2x_2 + \ldots + k_mx_m$, where ε is the error term, and $k_1, \ldots, k_m$ are the regression coefficients used to minimize the sum of squared errors over a set of training examples. Linear regression is appropriate if the relationship among variables is linear; however, most interesting real-world domains exhibit some degree of nonlinearity, which makes the modeling significantly more difficult. A decision tree with a linear regression model in each leaf can also approximate a nonlinear function. The idea with model trees [*Quinlan,* 1992; *Wang and Witten*, 1997] is that the problem of learning the possible nonlinear relationship between the dependent variable and the independent variables can be divided into n smaller subproblems of separately learning each dimension, or component. To separate a problem into n subproblems, a condition is needed, represented by an internal node tree. Each internal node in the tree contains a splitting decision based on the input variables $x_1, \ldots, x_m$ that divides the data into two subsets corresponding to the left and right sub-trees. Model trees, constructed by the algorithm called M5 [*Quinlan,* 1992; *Wang and Witten*, 1997], can have multivariate linear models in the leaves, as shown in the bottom left of Figure 3. M5 learns efficiently and can tackle tasks with very high dimensionality (up to hundreds of attributes).

Hansen and Salamon [1990] showed that the generalization ability of a model can be significantly improved by combining a number of models as an ensemble. Because of their simple and effective properties, ensembles have become a hot topic in the machine learning community. In ensemble modeling, multiple learning algorithms are used to obtain better predictive performance than could be obtained independently from any of the constituent learners. Empirically, ensembles tend to yield better results when there is significant diversity among the models; for that reason, many ensemble methods seek to promote diversity among the models they combine [*Kuncheva and Whitaker,* 2003].

There are several ensemble techniques that could be used to predict time series (i.e., the integral proton flux) from past temporal data. *Deng et al*. [2005] used support vector machine (SVM) as the ensemble component. *Wichard and Ogorzałek* [2007] proposed an ensemble constructed by several model classes, including ANN, Nearest-Neighbor Models. Instead of using artificial neural networks (ANN), nearest-neighbor models or the SVM as ensemble components, we used model trees [*Quinlan,* 1992; *Wang and Witten*, 1997], which are trees of regression models and which have been shown to be very accurate. Another difference from other ensemble approaches is that our strategy for promoting diversity within the ensemble is to construct model trees that predict different future values, which can be combined (through interpolations, extrapolations and weighted averaging) to produce the final prediction of a single value. In our approach, every model tree is constructed from past values of several time series to predict a single future value of one of the analyzed time series.

2.2.2. An ensemble approach using model trees for predicting poorly-connected SEP events

Our model trees are composed of temporal nodes, which are conditions about past values of proton flux time series; arcs, which are decision routes (yes/no); and temporal leaves, which contain multiple linear regression models that predict a value at some time in the future. For example, suppose that a node labeled "$logP7(t) > 0.5$" has two arcs ("yes" and "no") and that the "yes" arc is connected with the leaf "$logIpf(t+1) = 0.1 + 2\cdot logP7(t-1)$", as shown in the model tree MT_1 in the bottom left of Figure 3. The meaning of this regression rule, which includes the condition node as the antecedent and the leaf as the consequent, is: if a condition (based on the current value of *logP7*) is fulfilled, then a future value of *logIpf* may be estimated as a linear function of the past value of *logP7*. More specifically, it means that "if the base-10 logarithm of the value of the differential proton channel P7 at the current time t is greater than 0.5, then the base-10 logarithm of the integral proton flux in the next hour may be predicted using a linear function of the a past value of the base-10 logarithm of the P7 channel 1 hour ago". This rule may be understood because it is short; however, longer rules with tens of condition nodes that end in linear functions with tens of variables are barely understandable. Moreover, the prediction ensemble that we constructed, composed of 24 model trees in which each model tree is composed of hundreds of nodes and leaves, is not useful as a knowledge source but as a predictor. As is true for other "black box" data mining models (e.g., neural nets), the only purpose of our regression ensemble is to predict future values, and it is not possible to acquire understandable knowledge from it (i.e. relationships between the integral proton and the differential proton fluxes).

As we stated before, to construct the ensemble model, 24 model trees are trained from past data. The first model tree MT_1 forecasts the base-10 logarithm of the integral proton flux (E>10 MeV) at 1 hour ahead of the current time t, which we call $logIpf(t+1)$; the second model tree, MT_2, forecasts the same variable at 2 hours in the future, that is, $logIpf(t+2)$, and so on. The ensemble modeling and its use in the prediction of poorly-connected SEP events are illustrated in Figure 3.

The left-hand side of Figure 3 shows the offline learning process, which is done only once. Every model tree is constructed from a *training table* of examples. Each row of the table, a training example, is constructed every 5 minutes within a *training time interval*. We selected 202 *training time intervals* of several tens of hours each. The analyzed intervals were selected from situations where a gradual integral proton flux surpassed the SWPC SEP threshold (e.g., from 12:00 UT on December 5 to 16:00 UT on December 6, 2003) and situations where the gradual integral proton flux ended up decreasing to background levels (e.g. from 21:00 UT on July 29 to 21:00 UT on August 2, 2004). Because there are twelve 5-min time steps in an hour, an interesting historic situation of n hours of duration is described with $n\times12$ examples. Each example is a summary of a specific instant within a historic situation, in terms of its current and (recent) past values. Table 1 shows the fields of the *training table* of examples needed to construct a model tree. This *training table* has 161 fields, 160 for the input temporal variables and 1 field for the response variable. The response variable is what the model tree has to learn to predict. The input temporal variables give a snapshot of the recent past of the time when the record was constructed. Each field of a learning example is filled with a past value of one of the analyzed time series: $logP3$, $logP4$, $logP5$, $logP6$ and $\log P7$, which are the base-10 logarithm of the differential proton flux time series associated to energies from 9 MeV to 500 MeV; *DlogP3, DlogP4, DlogP5, DlogP6 and DlogP7*, which are the time series of the first derivatives of the corresponding proton channels; *logIpf* and *DlogIpf* are the base-10 logarithm of the integral proton flux (E>10 MeV) and its first derivatives, respectively; and, finally, calculated values from past values of the same input variables, such as the average of the last two or three derivatives of a single proton channel, which shows the log-linear behavior of a specific differential proton channel.

2.2.3. Issuing preliminary forecasts of poorly-connected SEP events

While running in real-time, each Model Tree checks the conditions of its nodes and responds with a prediction vote about a future value. Because there are 24 model trees, a vector of 24 prediction votes is obtained every time (the prediction vote for $logIpf(t+1)$, the prediction vote for $logIpf(t+2)$, and so on). In order to calculate the final predicted value at a single time in the future, say $logIpf(t+5)$, the corresponding prediction vote for $logIpf(t+5)$ is important, but the influences of the neighboring prediction votes are also important, depending on their temporal proximity. More specifically, the prediction of a value at $t+i$ is the polynomial interpolation (of order 3) of the predicted votes of the model trees MT_j, such that $j<i$, and the polynomial extrapolation (of order 3) of prediction votes of MT_k where $k>i$. The weights were adjusted empirically by maximizing the general performance precision (1-FAR) and recall (POD), as summarized at the end of Section 2.1.2, while predicting poorly-connected SEPs for solar cycles 22 and 23 with valid configurations of weights. A valid weight configuration is one in which the weight of a prediction vote j is inversely proportional to its distance to the value to be predicted, j-i, and the sum of weights is 1.

The expected time series of the integral proton flux is the sequence of final predicted values. The final predicted time series is finally used as input to make the poorly-connected SEP forecast: the expected time of surpassing 10 pfu, that is t_{10pfu}, and the maximum intensity seven hours later (from t_{10pfu} to t_{10pfu}+7 hours). The seven-hour limit was set simply to homogenize the outputs of both well-connected and poorly-connected SEP forecasting models. Poorly-connected SEP forecasts are sent to the high-level module, which has to either confirm them or filter them out.

2.3. Description of the high-level module for analysis and inference

Forecasting SEPs (E>10 MeV) may be seen as the synergy between well- and poorly-connected SEP forecasting models. The Analysis and Inference module receives the preliminary warnings from the well- and poorly-connected SEP forecasting models and checks consistency with all the received information. If a preliminary forecast is accepted as final, this module retrieves solar data with the purpose of calculating the maximum intensity of the first seven hours of the predicted proton event and (in the case of well-connected events) identifying the active region.

2.3.1. Consistency checking of preliminary warnings

This task is performed by applying empirical rules of consistency based on previous experiences and observations. The following are some of the empirical rules that are applied to filter out preliminary warnings:

- A prediction issued by the well-connected SEP forecasting model is filtered out when the time of the associated flare is farther than t-L-τ_1, where t is the current time, L is the transit time found for protons from the Sun to 1 AU and τ_1 is a threshold that needs to be defined empirically.
- In the case of poorly-connected SEPs, if the model predicts an event but the integral proton flux has been decreasing during τ_2 hours, the preliminary warning is filtered out. The value of τ_2 is another threshold that was defined empirically.
- If both forecasting models issue a prediction at the same time, this situation is commented to the user; however, the prediction details of the poorly-connected forecasted event are ignored because the details of the forecasted well-connected event are produced by a method that is based on evidence rather than on similarity to historic cases.
- If a preliminary forecast has been filtered out or if neither model issues any preliminary forecast, the Analysis and Inference module finally concludes that no SEP event is expected. It is important to mention neither forecasting models issue "non-SEP" forecasts; they only issue preliminary forecasts of SEP events.

The best values for the thresholds (i.e., τ_1 and τ_2) and rules of this module were identified empirically by maximizing the general forecasting performance measure, which includes precision (1-FAR) and recall (POD) as summarized at the end of Section 2.1.2, while predicting well-connected and poorly-connected SEPs for solar cycles 22 and 23.

Finally, we obtained several *optimal* configurations, among which it was not easy to decide which one was best. For example, for experiment 1 of Section 3, in which we calibrated the system with the solar data from solar cycle 22 to predict events from solar cycle 23, two threshold configurations were finally obtained:

Configuration I:

Probability of detection of SEPs:	80.65% (75//93
False alarm rate:	25.74% (26/101)
Average warning time:	4 h 31 min

Configuration II:

Probability of detection of SEPs:	86.02% (80/93)
False alarm rate:	32.77% (39/119)
Average warning time:	5 h 21 min

Configuration I, the most conservative, yields a worse POD, but it has a much better FAR than configuration II. The notably better FAR value represents a slightly better general performance. In experiment 2 of Section 3, for configuring the official UMASEP forecaster, we encountered a similar situation regarding one configuration with a better POD and another more conservative configuration with a similar POD but a much better FAR. In both experiments, the most conservative forecasts (with better FARs) were considered the final configurations.

2.3.2. Calculating the intensity of the first hours of the predicted proton event

It is very difficult to make real-time predictions of the peak intensity of an SEP event. The maximum intensity could occur several hours or even days after the official SEP onset time, and it depends on several phenomena. In complex cases, some of these phenomena occur after the integral proton flux has surpassed 10 pfu, such as additional flares, CMEs, and CIRs. However, a user needs to know the intensity of the first hours of the SEP event, which is possible with an acceptable error and depends on the type of expected event (well- or poorly-connected). In the case of poorly-connected SEP events, the forecast of the intensity of the integral proton flux (E>10 MeV) during the first hours of the expected SEP event is an output of the regression ensemble and posterior intra/extrapolations (Section 2.2.3) so the Analysis and Inference Module does not alter these estimations. However, the well-connected forecasting model does not need to predict integral proton fluxes for issuing the event onset forecast, so the prediction of the intensity of first hours has a different strategy — explained in the next two paragraphs—, and it is performed by the Analysis and Inference Module only if the preliminary forecast is not filtered out.

In the case of well-connected SEPs, our assumption is that the intensity of the prompt component is related to the magnetic connectivity and to the associated flare. We explored several functions to predict the intensity, say I, of the prompt component as a function of a variable M related to the magnetic connectivity and another variable F related to the flare, that is, $I=function_1(M, F)$. We assumed M to be a function of the magnetic connectivity; that is, $M=function_2(mc)$, where mc is the magnetic connectivity calculated as the maximum fluctuation similarity found at the moment in which a prediction is issued (see Section 2.1.2 for more details about how to calculate this lag-correlation measure). On the other hand, F is a function of the evolution of the X-ray flux, that is, $F=function_3(xrays)$, where *xrays* is the time series of the soft X-ray flux of the associated flare found (see Section 2.1.2 for more details). The next paragraph explains the empirical assumptions to encounter the functions (and their internal parameters) $function_1$, $function_2$, and $function_3$ that reduce the root mean squared error between I and the real value of the intensity of the prompt component.

We did not know if the functions were linear or nonlinear. If they were nonlinear, we had to find a satisfactory type of nonlinear functions. Therefore, we explored several possible functions, their internal parameters, and other parameters (e.g., the best temporal length of the prompt component, which finally was 7 hours) by trial and error tests that allowed a reduction in the intensity error of the prompt component for solar cycles 22 and 23. We found that a satisfactory function to calculate $I=function_1(M, F)$ was the linear function of the product M times F, that is $I = mX+b$, where m and b are the parameters of the linear function $X= M{\cdot}F$, and M and F are the functions of the magnetic connectivity and the associated flare, respectively. To find a good function $F=function_3(xrays)$, we explored other functions on the soft X-ray flux that have been used in SEP forecasting. It is well established that the time-integrated soft X-ray flux is related to SEP events [*Kubo and Akioka*, 2004]. *Balch* [2008] and *Laurenza et al,* [2009] used this approach to predict the occurrence of an SEP event. We used a similar approach, the first half of the time-integrated *xrays,* for calculating F. In other words, if t_1 is the time of the maximum intensity of the associated flare, say r, and t_2 is the previous time when the intensity was greater than a percentage p of r, we calculated F as the time-integrated *xrays* from t_1 to t_2. We found that a satisfactory function to calculate $M=function_2(mc)$ was the exponential function $M=10^{mc}$.

The final model output includes intensity 'bands' of varying thickness, showing the uncertainty of the SEP predictions. The uncertainty of the predicted SEP start time (time of meeting or surpassing the SWPC SEP threshold) is presented as $[t+t_{min}, t+t_{max}]$, and the uncertainty of the predicted intensity I is presented as $I \pm \delta I$, where t is the current time (that is, the time in which the forecast is issued), t_{min} and t_{max} are the extremes of the expected time interval of occurrence, I is the intensity value predicted by the forecasting model and δ is a thickness percentage. Any SEP forecast has the form $\{[t+t_{min}, t+t_{max}], I \pm \delta I\}$, which means that an SEP event is expected to occur after $t + t_{min}$ and before $t + t_{max}$, with a minimum intensity of $I - \delta I$ and maximum intensity of $I + \delta I$. The estimation of t_{min}, t_{max} and δ depends on the type of event and includes some parameters that were estimated for the purpose of maximizing the POD and minimizing the FAR using data from solar cycles 22 and 23. In the case of well-connected SEP event forecasting, $t_{min} = 0$, t_{max} is the minimum value between 2 hours and T, and δ=0.23, where T is the transit time of protons (see Section 2.1.2). In the case of poorly connected SEP event predictions $t_{min} = 30$ minutes, t_{max} is the forecasted value of the ensemble of regression modes, and δ=0.02.

2.3.3. Identifying associated flares and active regions of well-connected events

When an event has been predicted by the well-connected SEP forecasting model, the method presented in Section 2.1 identifies the specific time interval of the associated soft X-ray fluxes. If the forecast is not filtered out, the Analysis and Inference module retrieves the NOAA solar event List file (http://www.swpc.noaa.gov/ftpdir/indices/events/events.txt) with details about recent solar events. This file is collected with the sole purpose of providing useful information about the situation to the user, specifically, the active region and/or heliolongitude of the associated flare. Because the NOAA solar event list is updated every 30 minutes, some well-connected SEP predictions may be shown without the corresponding region and/or heliolongitude.

UMASEP does not use the NOAA solar event list to make any forecasting analysis or any decision (e.g., magnetic connection estimation, forecasting, filtering). The only purpose of reading the NOAA solar event List (when it is available) is to show complementary information to the space weather user.

When there is a poorly-connected SEP forecast, the Analysis and Inference module does not consult the

Event List file; therefore, no associated flare, region or heliolongitude is identified.

3. Outputs of the UMASEP system

The graphical output of UMASEP is updated automatically in the forecast panel (http://spaceweather.uma.es/forecastpanel.htm). Figure 4 shows the forecast panel that an operator would have seen if UMASEP had processed the real-time GOES data on October 26, 2003. This figure also shows inferences about the associated flare, heliolongitude and active region, as well as a small illustration of a possible route of the solar protons from the corresponding heliolongitude toward the near-Earth environment.

The upper time series shows the integral proton flux with energies greater than 10 MeV. The current flux is indicated below the label "now". To the right of this label, the forecasted integral proton flux is presented. Colors indicate the intensity of the expected integral proton flux at that specific time.

The middle time series shows recent solar activity in terms of soft X-rays, and the lower time series shows the magnetic connectivity with the most well-connected CME/flare process zone. When a forecast is issued, the graphical output also shows the details of these predictions and what the model infers about the situation. Figure 4 shows the output of the forecaster after processing the data from October 2003. At 18:00 UT on

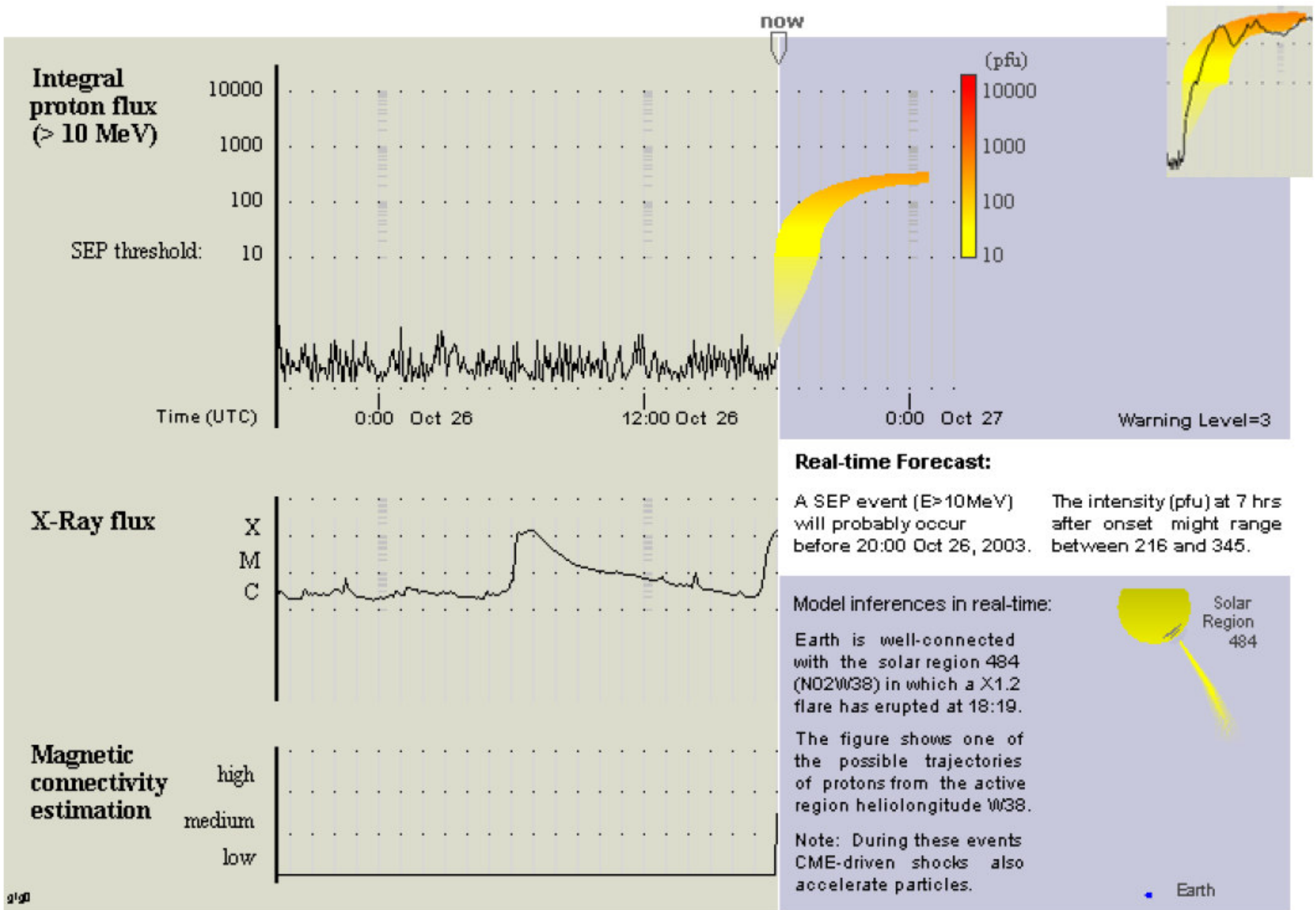

Figure 4. This figure shows the UMASEP output after processing GOES-10 data from October 26, 2003. The small upper-right chart is not part of the forecaster output; it shows the posterior evolution of the integral proton flux for this event, the first "Halloween" SEP event, showing that the forecast was successful. Note that the well-connected SEP was forecasted when no enhancement in the integral proton flux (E> 10 MeV) was observed (see also Figure 2a) because the correlated rise of proton channel P6 (see Figure 2a) occurred with low flux (as well as the rest of the differential proton channels).

October 26, the integral flux can be seen to be almost flat, but the forecaster issues a prediction. The forecast details showed that the expected event would arrive during the following two hours and that it would reach an intensity in the range of two hundred to three hundred pfu. Below the forecast section, the system also presents the model inference section, which shows that the Earth is well-connected with the solar region 484, in which a solar flare has erupted. Because this flare is historic, the system shows the historic peak time and intensity. The system also shows that the associated heliolongitude is west 38. On the lower right section of the forecast panel, the system also presents a graphical illustration of the Sun-Earth link, showing a possible trajectory of the predicted protons. This illustration is useful to show that no solar proton event has arrived at Earth but that one is coming along the magnetic field lines. As time passes, the integral proton flux also rises. The system also refines the forecasted intensity to the band between two hundred and six hundred pfu. At 18:25 UT, the flux has surpassed the 10-pfu threshold, which indicates that an official proton event is occurring. The forecast was successful with a warning time of 25 minutes, as is shown in the small image at the top right section of Figure 4.

Figures 5 and 6 show the different forecast strategies of the well-connected and the poorly-connected models. Figure 5 shows two successful forecasts of well-connected SEP events that occurred on November 8, 1987, and April 21, 2002. The warning times were 3 h 45 min and 35 minutes respectively. The inferences of the last event are shown (the associated flare and active region, as well as an illustration of one of the possible paths of solar protons). Figure 6 shows the forecast of two poorly-connected SEPs during December 7, 2006 (upper chart), and April 16, 1990 (lower chart). The respective warning times were 20 h 05 min and 22 h 05 min. Note that the left-hand chart of Figure 5 shows a well-connected SEP prediction that is very satisfactory: the prediction was issued almost from background (low) flux levels, had a good warning time, and predicted the integral proton flux seven hours after the flux surpassed 10 pfu.

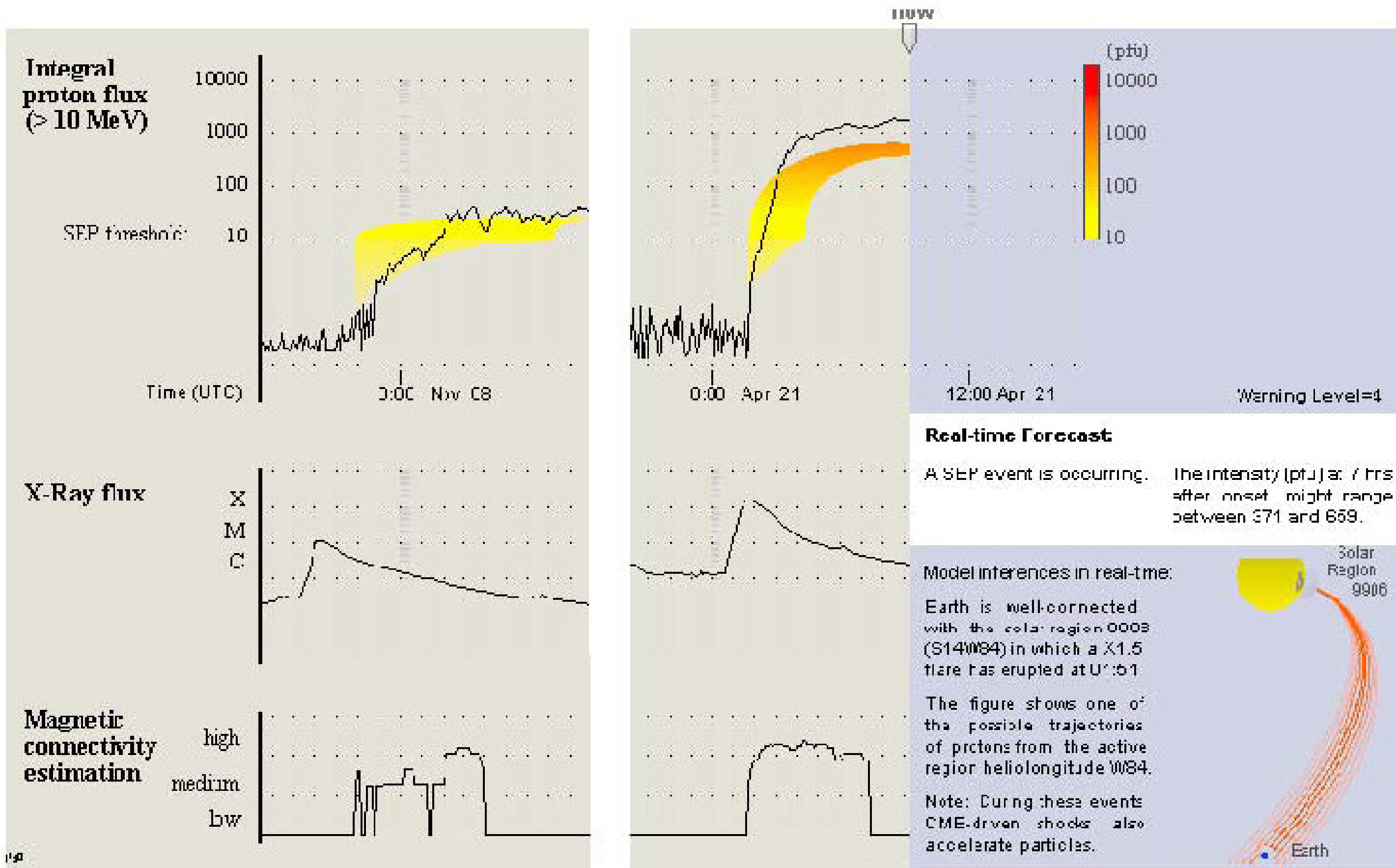

Figure 5. Several successful forecasts of well-connected SEP events that occurred on 8 November 1987 and 21 April 2002. The inferences of the last event are shown (the associated flare and active region, as well as an illustration of one of the possible paths of solar protons).

The right-hand chart in Figure 5 shows a successful well-connected SEP forecast, but the quality of the result is not satisfactory, perhaps owing to the fast nature of the phenomenon: the well-connected forecast was issued from the middle of the rise of the integral proton flux, the warning time was low, and the forecasted intensity of the prompt component had a high error, nearly 0.6, in terms of the base-10 logarithm of the integral proton flux. On the other hand, Figure 6 shows a different strategy: the poorly-connected SEP forecasting model needs a flux rise in several differential proton channels (and therefore a rise of integral proton flux) to issue a prediction. In poorly-connected SEP forecasts, there is no information about the associated active region and flare. There is also an animation that shows UMASEP in action while predicting several SEP events . Section A.4 of the Auxiliary Material presents more UMASEP`s results after analyzing solar data of the Sun-Earth link situations during solar 22 and 23. The video shows the output of the system when situations of successful predictions, missed events and false alarms take place.

4. Statistical performance

The most common metrics for measuring the performance of SEP event predictors are the POD, FAR and warning time (WT). These metrics have been widely used in recent papers and presentations about automatic SEP forecasters [*Balch,* 2008; *Laurenza et al.*, 2009; *Posner,* 2007].

SEP event forecasting performance measures use the following variables: number of correct forecasts or hits, A (an SWPC SEP event was forecasted and one occurred); the number of false alarms, B (an SWPC SEP event was forecasted but none occurred); the number of missed events, C (no SWPC SEP event was predicted but an event did occur); the number of correct nulls, D (no SWPC SEP event was forecasted and none occurred). Then, POD = A/(A + C) and FAR = B/(A + B).

It is useful to provide the performance in predicting very

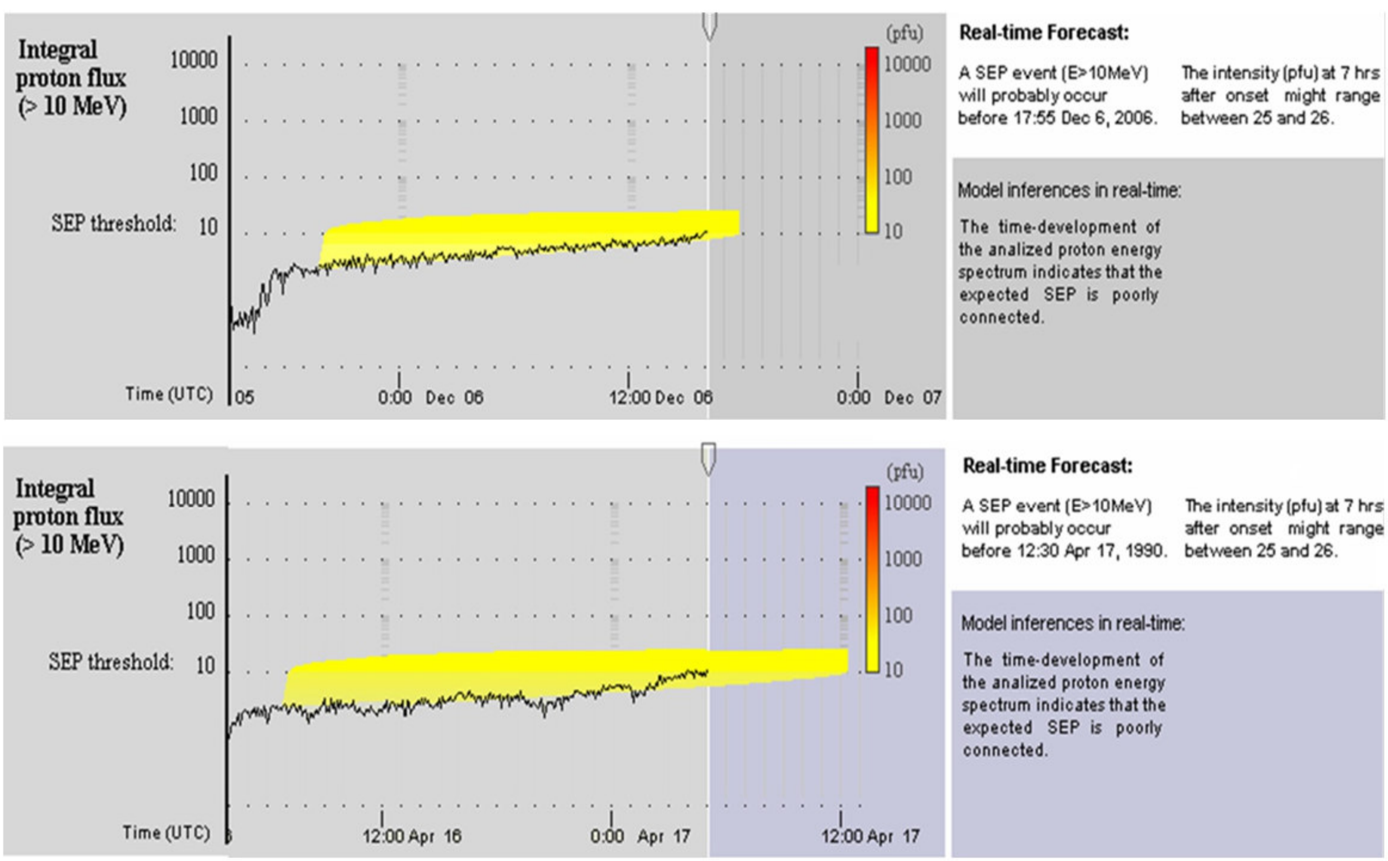

Figure 6. This figure shows several successful forecasts of poorly-connected SEP events. This figure shows the forecast of two SEPs during December 7th, 2006 (upper chart) and April 16th, 1990 (lower chart). The respective warning times were 20 h 05 min and 22 h 05 min.

fast SEPs, also called prompt SEPs, which are detected near Earth in eight hours or less after the associated major solar event. The rest of the events, the delayed SEPs, are those events that are detected more than eight hours after the major solar event or that are not associated with a flare. The lapse of eight hours is the time that results in the same number of prompt and delayed events that have occurred since 1987 (solar cycles 22 and 23).

Table 2 presents the list of events from the SWPC SEP list and the forecast results for each SEP event of solar cycles 22 and 23. From left to right the columns show the following:

- Event number.
- Start times (*ST*) of SEP events, are presented according to the SWPC table: http://www.swpc.noaa.gov/ftpdir/indices/SPE.txt . We detected two discrepancies in the SEP start times based on the integral proton fluxes (E>10 MeV) of the available GOES satellites: For the event 85 and 133, we changed the original start times (18:04 UT and 17:75 UT, respectively) to 18:40 UT and 17:55 UT, respectively. These corrections did not affect the POD and FAR.
- Integral proton flux (E>10 MeV) at *ST* + 7 hours, where *ST* is the start time of the SEP event.
- Type of effect on the proton enhancement before the *ST*: prompt and delayed events.
- Forecast results: *Hits* are those SEP events forecasted with a warning time greater than or equal to one minute. *Misses* are those events that were not anticipated.
- The warning time, which is the temporal difference between the start time of the SEP event, *ST*, and the time at which the forecast was issued.
- Root mean squared error of the predicted integral proton flux (E>10 MeV) at *ST* + 7 hours. The error was calculated between the log10 of the real value of the integral proton flux at *ST* + 7 hours and the average of the predicted band of values (minimum and maximum, in terms of log10).

The statistical performance from two evaluation experiments is presented and summarized in Table 3: Table 3a shows the performance of a version of UMASEP whose thresholds, parameters and rules were adjusted for cycle 22 and whose evaluation was performed with the data from solar cycle 23. Table 3b shows the evaluation of the official and online version of UMASEP, whose thresholds, parameters and rules were adjusted for both solar cycles 22 and 23 and whose evaluation was performed with the data from the solar cycles 22 and 23. The purpose of the first experiment was to evaluate the generalization capability of the model by predicting unseen cases; the second experiment evaluates the performance of the final system with the best possible tuning to face future cases.

For the first experiment, the performance is shown in Table 3a by using the following: the well-connected SEP forecasting model and the Inference and Analysis module (abbreviated as the WC-model), the poorly-connected SEP forecasting model and the Inference and Analysis module (abbreviated as the PC-model), and both models and the Analysis and Inference module (abbreviated as the WC-PC-model, which is the UMASEP).

In Table 3a and 3b, each model (WC and PC) is applied over the entire time interval (solar cycles 22 and 23). Because they are different models, they issue different predictions (i.e., the WC-model issues different predictions than PC-model for the same time interval); consequently, they issue a different number of false alarms. The denominator of the false alarm rate is the number of positive forecasts, and therefore, the FAR of each model is different.

Tables 4a and 4b show the performance of UMASEP for different sizes of SEP events and different levels of solar activity. Table 4a shows contradictory performance measures for the strongest events: it had the best performance in terms of the POD for severe and extreme SEP events but the worst estimation of the intensity seven hours after the start of these events. For the rest of the SEP categories, the POD and intensity error are similar to the average UMASEP performance. Table 4b shows the performance of UMASEP for different levels of solar activity. This table shows that during low solar activity, UMASEP performs better. We think that during low solar activity, the magnetic connections between the Earth and the proton acceleration source (e.g., parent solar events and interplanetary CME-driven shocks) are less affected by other strong solar phenomena (e.g., CME and CIRs); therefore, magnetic connections are more defined and less chaotic, allowing UMASEP to predict better the proton events.

Table 2. SWPC SEP list and the forecast result for each event of solar cycles 22 and 23.

Solar Cycle 22

Event Number	Start Time (*ST*) of SEP Event[a]	Proton Flux at *ST*+7hrs (pfu)[b]	Type of Solar Event[c]	SEP Forecast Result[d]	Warning Time[e]	RMS error of Flux at *ST*+7hrs (log pfu)[f]
1	11/08/1987 2:00	41	Prompt	Hit	4 h	0.12
2	01/02/1988 23:25	79	Prompt	Hit	10 min	0.24
3	03/25/1988 22:25	33	Prompt	Miss		
4	06/30/1988 10:55	13	Prompt	Miss		
5	08/26/1988 0:00	12	Delayed	Miss		
6	10/12/1988 9:20	7	Prompt	Hit	2 h 5 min	1.30
7	11/08/1988 22:25	7	Delayed	Hit	3 h 20 min	0.59
8	11/14/1988 1:30	7	Prompt	Hit	50 min	0.57
9	12/17/1988 6:10	9	Delayed	Miss		
10	12/17/1988 20:00	13	Delayed	Hit	10 h 55 min	0.30
11	01/04/1989 23:05	11	Prompt	Hit	1 h 5 min	0.65
12	03/08/1989 17:35	202	Delayed	Hit	12 h 40 min	0.81
13	03/17/1989 18:55	277	Prompt	Hit	5 min	0.66
14	03/23/1989 20:40	30	Prompt	Hit	20 min	1.13
15	04/11/1989 14:35	123	Delayed	Hit	10 h 30 min	0.68
16	05/05/1989 9:05	10	Delayed	Miss		
17	05/06/1989 2:35	45	Delayed	Hit	17 h 5 min	0.25
18	05/23/1989 11:35	68	Delayed	Hit	5 h 40 min	0.43
19	05/24/1989 7:30	8	Delayed	Hit	16 h 40 min	0.52
20	06/18/1989 16:50	11	Prompt	Hit	25 min	0.37
21	06/30/1989 6:55	6	Delayed	Hit	24 h	0.65
22	07/01/1989 6:55	6	Delayed	Hit	4 h 35 min	0.62
23	07/25/1989 9:00	32	Prompt	Miss		
24	08/12/1989 16:00	782	Prompt	Hit	5 min	0.11
25	09/04/1989 1:20	17	Delayed	Miss		
26	09/12/1989 19:35	19	Delayed	Hit	6 h 5 min	0.69
27	09/29/1989 12:05	1850	Prompt	Hit	10 min	0.18
28	10/06/1989 0:50	22	Delayed	Hit	45 min	0.06
29	10/19/1989 13:05	1530	Prompt	Miss		
30	11/09/1989 2:40	16	Delayed	Miss		
31	11/15/1989 7:35	38	Prompt	Hit	15 min	0.97
32	11/27/1989 20:00	38	Delayed	Hit	15 min	0.17
33	11/30/1989 13:45	256	Prompt	Miss		
34	03/19/1990 7:05	315	Prompt	Hit	35 min	0.05
35	03/29/1990 9:15	16	Delayed	Hit	11 h	0.21
36	04/07/1990 22:40	17	Delayed	Hit	10 h 10 min	0.19
37	04/11/1990 21:20	13	Delayed	Miss		
38	04/17/1990 5:00	12	Delayed	Hit	22 h 5 min	0.33
39	04/28/1990 10:05	119	Delayed	Hit	2 h 20 min	0.59
40	05/21/1990 23:55	400	Prompt	Hit	55 min	0.12
41	05/24/1990 21:25	177	Prompt	Hit	5 min	0.64
42	05/28/1990 7:15	31	Delayed	Miss		
43	06/12/1990 11:40	79	Prompt	Hit	2 h 55 min	0.09
44	07/26/1990 17:20	21	Delayed	Hit	14 h 35 min	0.15
45	08/01/1990 0:05	18	Delayed	Hit	1 h 45 min	0.15
46	01/31/1991 11:30	240	Delayed	Hit	3 h 35 min	0.19
47	02/25/1991 12:10	13	Prompt	Hit	35 min	1.10
48	03/23/1991 8:20	2260	Delayed	Hit	10 min	1.30
49	03/29/1991 21:20	20	Delayed	Miss		
50	04/03/1991 8:15	26	Delayed	Hit	5 h	0.00
51	05/13/1991 3:00	350	Prompt	Hit	40 min	0.19
52	05/31/1991 12:25	20	Delayed	Hit	2 h	0.11
53	06/04/1991 8:20	44	Prompt	Miss		
54	06/14/1991 23:40	18	Delayed	Miss		
55	06/30/1991 7:55	19	Delayed	Hit	7 h 35 min	0.18
56	07/07/1991 4:55	35	Prompt	Miss		
57	07/11/1991 2:40	30	Delayed	Miss		
58	07/11/1991 22:55	14	Delayed	Hit	20 h 5 min	0.26
59	08/26/1991 17:40	31	Delayed	Hit	16 h 5 min	0.09
60	10/01/1991 17:40	12	Delayed	Miss		

Solar Cycle 23

Event Number	Start Time (*ST*) of SEP Event[a]	Proton Flux at *ST*+7hrs (pfu)[b]	Type of Solar Event[c]	SEP Forecast Result[d]	Warning Time[e]	RMS error of Flux at *ST*+7hrs (*log* pfu)[f]
74	11/04/1997 8:30	72	Prompt	Hit	50 min	0.06
75	11/06/1997 13:05	310	Prompt	Hit	30 min	0.38
76	04/20/1998 14:00	361	Prompt	Hit	1 h 55 min	1.15
77	05/02/1998 14:20	149	Prompt	Hit	15 min	0.35
78	05/06/1998 8:45	210	Prompt	Hit	15 min	0.24
79	08/24/1998 23:55	111	Prompt	Hit	1 h	0.19
80	09/25/1998 0:10	44	Delayed	Hit	2 h 20 min	0.15
81	09/30/1998 15:20	931	Prompt	Hit	1 h	1.56
82	11/08/1998 2:45	11	Delayed	Hit	12 h 30min	0.38
83	11/14/1998 8:10	310	Prompt	Miss		
84	01/23/1999 11:05	14	Delayed	Hit	20 h 5 min	0.26
85	04/24/1999 18:40	32	Delayed	Miss		
86	05/05/1999 18:20	14	Delayed	Miss		
87	06/02/1999 2:45	48	Delayed	Hit	2 h	0.25
88	06/04/1999 9:25	64	Prompt	Hit	4 h 30 min	0.40
89	02/18/2000 11:30	13	Delayed	Miss		
90	04/04/2000 20:55	35	Prompt	Hit	2 h 55 min	0.14
91	06/07/2000 13:35	28	Delayed	Hit	4 h 55 min	0.04
92	06/10/2000 18:05	46	Prompt	Hit	30 min	0.04
93	07/14/2000 10:45	8000	Prompt[g]	Hit	24 h	1.09
94	07/22/2000 13:20	17	Prompt	Hit	45 min	0.33
95	07/28/2000 10:50	18	Delayed	Hit	6 h 40 min	0.28
96	08/11/2000 16:50	17	Delayed	Hit	20 min	0.37
97	09/12/2000 15:55	170	Prompt	Hit	1 h 5 min	0.82
98	10/16/2000 11:25	15	Prompt	Hit	3 h 10 min	0.28
99	10/26/2000 0:40	15	Delayed	Hit	8 h 45 min	0.33
100	11/08/2000 23:50	10700	Prompt	Hit	10 min	1.91
101	11/24/2000 15:20	94	Delayed	Hit	6 h 20 min	0.50
102	01/28/2001 20:25	33	Prompt	Hit	2 h 30 min	0.11
103	03/29/2001 16:35	27	Prompt	Hit	3 h 10 min	0.26
104	04/02/2001 23:40	488	Prompt	Hit	10 min	0.87
105	04/10/2001 8:50	65	Prompt[g]	Hit	15 h 35min	0.13
106	04/15/2001 14:10	951	Prompt	Hit	5 min	0.23
107	04/18/2001 3:15	271	Prompt	Hit	5 min	1.03
108	04/28/2001 4:30	57	Delayed	Hit	8 h	0.35
109	05/07/2001 19:15	28	Delayed	Hit	2 h 30 min	0.01
110	06/15/2001 17:50	26	Delayed	Miss		
111	08/10/2001 10:20	17	Delayed	Hit	9 h	0.18
112	08/16/2001 1:35	493	Delayed	Miss		
113	09/15/2001 14:35	11	Prompt	Hit	1 h 15 min	0.37
114	09/24/2001 12:15	1010	Prompt	Hit	30 min	0.26
115	10/01/2001 11:45	279	Prompt	Miss		
116	10/19/2001 22:25	11	Prompt[g]	Hit	18 h 5 min	0.37
117	10/22/2001 19:10	24	Prompt	Hit	40 min	0.79
118	11/04/2001 17:05	1870	Prompt	Hit	15 min	1.15
119	11/19/2001 12:30	17	Delayed	Hit	1 h 55 min	0.19
120	11/22/2001 23:20	1470	Prompt	Hit	30 min	1.69
121	12/26/2001 6:05	779	Prompt	Hit	5 min	0.70
122	12/29/2001 5:10	76	Delayed	Hit	18 h 30min	0.47
123	12/30/2001 2:45	14	Delayed	Miss		
124	01/10/2002 20:45	80	Delayed	Hit	7 h 5 min	0.49
125	01/15/2002 14:35	15	Delayed	Miss		
126	02/20/2002 7:30	13	Prompt	Hit	20 min	0.29
127	03/17/2002 8:20	13	Delayed	Hit	20 h 40min	0.29
128	03/18/2002 13:00	26	Delayed	Hit	2 h 55 min	0.00
129	03/20/2002 15:10	19	Delayed	Miss		
130	03/22/2002 20:20	13	Delayed	Hit	3 h 50 min	0.29
131	04/17/2002 15:30	24	Prompt	Hit	1 h 55 min	0.09
132	04/21/2002 2:25	1800	Prompt	Hit	35 min	0.56
133	05/22/2002 17:55	61	Delayed	Hit	7 h 5 min	0.31

Table 2. (continued)

Solar Cycle 22

Event Number	Start Time (*ST*) of SEP Event[a]	Proton Flux at *ST*+7hrs (pfu)[b]	Type of Solar Event[c]	SEP Forecast Result[d]	Warning Time[e]	RMS error of Flux at *ST*+7hrs (log pfu)[f]
61	10/28/1991 13:00	40	Delayed	Hit	2 h 50 min	0.16
62	10/30/1991 7:45	94	Prompt	Hit	35 min	0.83
63	02/07/1992 6:45	78	Delayed	Hit	5 h 40 min	0.47
64	03/16/1992 8:40	10	Delayed	Hit	24 h	0.90
65	05/09/1992 10:05	55	Delayed	Hit	3 h 25 min	0.22
66	06/25/1992 20:45	244	Prompt	Hit	15 min	0.40
67	08/06/1992 11:45	14	Delayed	Hit	2 h 20 min	0.29
68	10/30/1992 19:20	1550	Prompt	Hit	30 min	0.26
69	03/04/1993 15:05	17	Prompt	Hit	1 h 35 min	0.18
70	03/12/1993 20:10	44	Prompt	Hit	1 h 20 min	0.30
71	02/20/1994 3:00	74	Prompt	Hit	50 min	0.06
72	10/20/1994 0:30	35	Prompt	Hit	2 h 15 min	0.01
73	10/20/1995 8:25	63	Prompt	Hit	45 min	0.39

Solar Cycle 23

Event Number	Start Time (*ST*) of SEP Event[a]	Proton Flux at *ST*+7hrs (pfu)[b]	Type of Solar Event[c]	SEP Forecast Result[d]	Warning Time[e]	RMS error of Flux at *ST*+7hrs (log pfu)[f]
134	07/07/2002 18:30	22	Prompt	Hit	4 h 55 min	0.06
135	07/16/2002 17:50	45	Delayed	Hit	2 h 25 min	0.16
136	07/19/2002 10:50	13	Delayed	Miss		
137	07/22/2002 6:55	24	Delayed	Hit	12 h 5 min	0.02
138	08/14/2002 9:00	20	Prompt	Hit	1 h 35 min	0.16
139	08/22/2002 4:40	36	Prompt	Hit	1 h 15 min	0.03
140	08/24/2002 1:40	317	Prompt	Hit	5 min	0.24
141	09/07/2002 4:40	28	Delayed	Hit	19 h 40 min	0.04
142	11/09/2002 19:20	197	Prompt	Hit	2 h 10 min	0.70
143	05/28/2003 23:35	15	Delayed	Hit	11 h 40 min	0.24
144	05/31/2003 4:40	27	Prompt	Hit	1 h 35 min	0.63
145	06/18/2003 20:50	16	Delayed	Hit	4 h 45 min	0.20
146	10/26/2003 18:25	466	Prompt	Hit	25 min	0.10
147	10/28/2003 12:15	5780	Prompt	Hit	35 min	0.23
148	11/02/2003 11:05	50	Delayed	Miss		
149	11/04/2003 22:25	200	Prompt	Miss		
150	11/21/2003 23:55	13	Delayed	Hit	1 h 50 min	0.29
151	12/02/2003 15:05	86	Prompt	Hit	10 min	0.53
152	04/11/2004 11:35	28	Prompt	Hit	3 h 35 min	0.05
153	07/25/2004 18:55	55	Prompt	Hit	1 h 30 min	0.33
154	09/13/2004 21:05	273	Delayed	Hit	3 h	1.03
155	09/19/2004 19:25	57	Prompt	Hit	50 min	0.35
156	11/01/2004 6:55	63	Delayed	Miss		
157	11/07/2004 19:10	495	Prompt	Hit	3 h 40 min	0.69
158	01/16/2005 2:10	79	Prompt[g]	Hit	15 h 20 min	0.44
159	05/14/2005 5:25	76	Delayed	Hit	3 h 15 min	0.47
160	06/16/2005 22:00	44	Prompt	Hit	20 min	0.08
161	07/14/2005 2:45	13	Delayed	Hit	4 h 5 min	0.32
162	07/27/2005 23:00	31	Delayed	Hit	18 h	0.09
163	08/22/2005 20:40	317	Prompt[g]	Hit	17 h 30 min	0.37
164	09/08/2005 2:15	37	Delayed	Hit	2 h	1.13
165	12/06/2006 15:55	28	Delayed	Hit	20 h 5 min	0.04
166	12/13/2006 3:10	698	Prompt	Hit	10 min	0.05

[a] Start times (*ST*) of SEP events are presented according to the NOAA/SWPC SEP event list (http://www.swpc.noaa.gov/ftpdir/indices/SPE.txt). Times are presented in UT. NOAA defines the start of a proton event to be the first of three consecutive data points with fluxes (E > 10 MeV) greater than or equal to 10 pfu. We detected two discrepancies in the SEP start times based on the integral proton fluxes (E>10 MeV) of the available GOES satellites. For the event 85 and 133, we changed the original start times (18:04 UT and 17:75 UT, respectively) to 18:40 UT and 17:55 UT, respectively. These corrections did not affect the POD and FAR.

[b] Integral proton flux (E>10 MeV) at *ST* + 7 hours, where *ST* is the start time of the SEP event.

[c] Type of effect on the proton enhancement before *ST* due to different conditions between the solar parent event and the Earth. For evaluation purposes, we classified the SEPs in terms of their type of effects on proton enhancement: prompt and delayed. Prompt SEPs are those events detected near Earth within eight hours after the associated major solar event. The rest of the events, the delayed SEPs, are those events that are detected in more than eight hours after the major solar event or are not associated with a flare according to the SWPC SEP event list.

[d] *Hits* are those SEP events forecasted with a warning time greater than or equal to one minute. *Misses* are those events that were not anticipated.

[e] The warning time is the temporal difference between the start time of the SEP event, *ST*, and the time at which the forecast was issued.

[f] Root mean squared error of the predicted integral proton flux (E>10 MeV) at the SEP *ST* + 7 hours. The error was calculated between the $\log_{10}$ of the real value of the integral proton flux at *ST* + 7 hours and the average of the predicted band of values (min and max, in terms of $\log_{10}$).

[g] The SEP event, classified as prompt for having been associated with a recent flare, was predicted much earlier by the poorly-connected forecasting model, and that is why the warning time is large. It is possible that the successful forecast was due to a fortuitous situation favorable to the forecaster; it is also possible that the event was poorly-connected and therefore that the observed flare occurrence was coincidental; another possibility is that the integral proton flux at 1 AU was the sum of poorly- and well-connected proton fluxes.

Table 3. Evaluation of two versions of UMASEP using data from solar cycles 22 and 23.

a) Experiment 1. Performance using data from solar cycle 23. Parameter adjustments using data from solar cycle 22.

	Probability of detection (POD)			False alarm rate (FAR)	Average warning time	RMS error (based on $\log_{10}$ values)
	prompt SEPs	delayed SEPs	all SEPs			
WC-model (only)	84.31% (43/51)	4.76% (2/42)	48.39% (45/93)	21.05% (12/57)	1 h 8 min	0.521
PC-model (only)	17.65% (9/51)	66.67% (28/42)	39.78% (37/93)	32.73% (18/55)	7 h 49 min	0.409
WC and PC models	92.16% (47/51)	66.67% (28/42)	80.65% (75/93)	25.74% (26/101)	4 h 31 min	0.474

b) Experiment 2. Performance using solar cycles 22 and 23 (official version of UMASEP).

	Probability of detection (POD)			False alarm rate (FAR)	Average warning time (hrs)	RMS error (based on $\log_{10}$ values)
	prompt SEPs	delayed SEPs	all SEPs			
WC-model (only)	81.93% (68/83)	6.02% (5/83)	43.98% (73/166)	30.48% (32/105)	1 h 5 min	0.472
PC-model (only)	16.87% (14/83)	72.29% (60/83)	44.58% (74/166)	41.73% (53/127)	8 h 28 min	0.366
WC and PC models	87.95% (73/83)	73.49% (61/83)	80.72% (134/166)	33.99% (69/203)	5 h 10 min	0.409

In statistics, a result is considered statistically significant if it is unlikely to have occurred by chance alone according to a pre-determined threshold probability, the significance level. For instance, if we calculate the POD by analyzing a small time interval with only five SEP events, our conclusion has low statistical significance. However, if we calculate the POD by analyzing a time interval with 1000 SEP events, our conclusion has a much higher statistical significance. Statistics should prove that the probability that our conclusion arises by chance is very low, that is, that the significance level α is very low (say 1% or 5%). Because the number of SEP events that we analyzed is not large (166), we have to prove that our conclusions regarding POD and FAR are statistically significant. Hypothesis tests are often used in science and social research to give support to conclusions; however, they have not been used in the field of SEP forecasting. The more SEP events that are taken into account during the validation, the more confidence we have in the POD and FAR. UMASEP has been evaluated with SWPC SEP events from two solar cycles, while the rest of the forecasters have been evaluated with a lower number of events. The next paragraph provides our hypothesis tests on POD and FAR and explains our calculations so that our system can be compared with other systems if statistical significance tests are used.

For experiments 1 and 2, we divided the analyzed time interval into continuous periods with equal numbers of contiguous SEP events. For the first experiment, we divided the evaluated solar cycle 23 into 7 time intervals, each with 14 contiguous SEP events, except for the last period of 9 SEPs. For the second experiment, we divided the evaluated time interval composed of solar cycles 22 and 23 (since September 1986) into 12 time intervals, each one with 14 contiguous SEP events, except for the last period of 10 SEPs. The duration of each tested continuous interval was different in all cases, but each interval included the same number of SEP events (with the exception of the last time interval). The statistical distribution of PODs and FARs calculated for each time interval did not follow a normal distribution; thus, we needed a nonparametric test for the hypotheses. We selected the Wilcoxon test [*Wilcoxon*, 1945], which allows the testing of the hypothesis that the averaged POD was greater than a satisfactory threshold and that the averaged FAR was lower than a satisfactory threshold. A indirect consequence of this test was the observation of how far the encountered thresholds were from the results of Tables 3a and 3b. The statistical significance of the experiments 1 and 2 was verified by applying a one-tailed Wilcoxon signed rank test with a significance level of 5%.

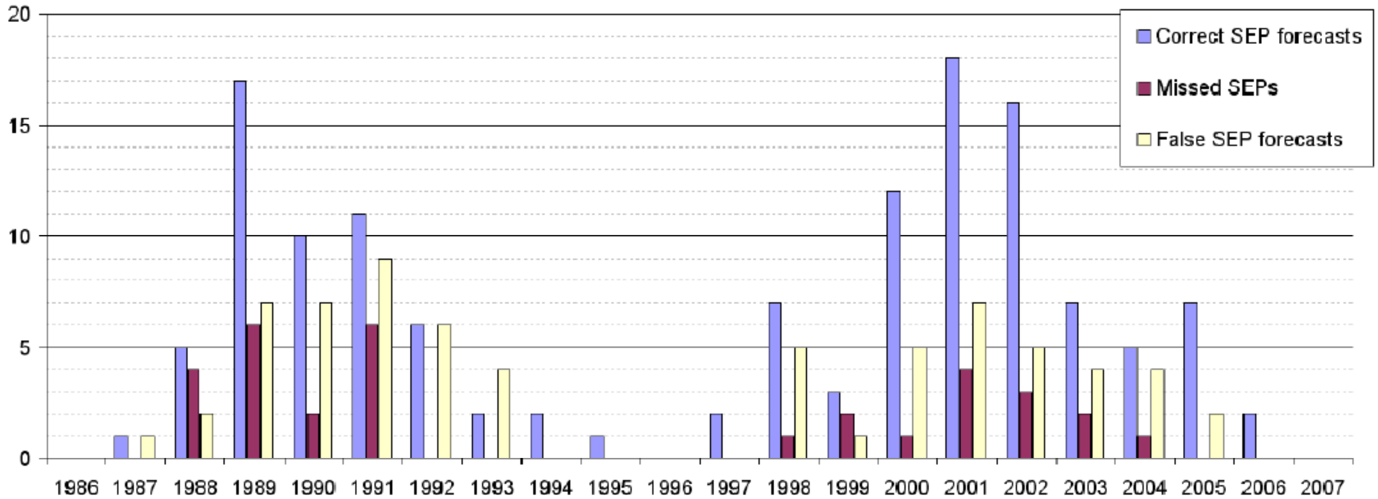

Figure 7. This figure shows the overall performance of UMASEP for every year of solar cycle 22 and 23 in terms of several counters: number of correct SEP forecast, number of missed SEPs and the number of false SEP forecasts.

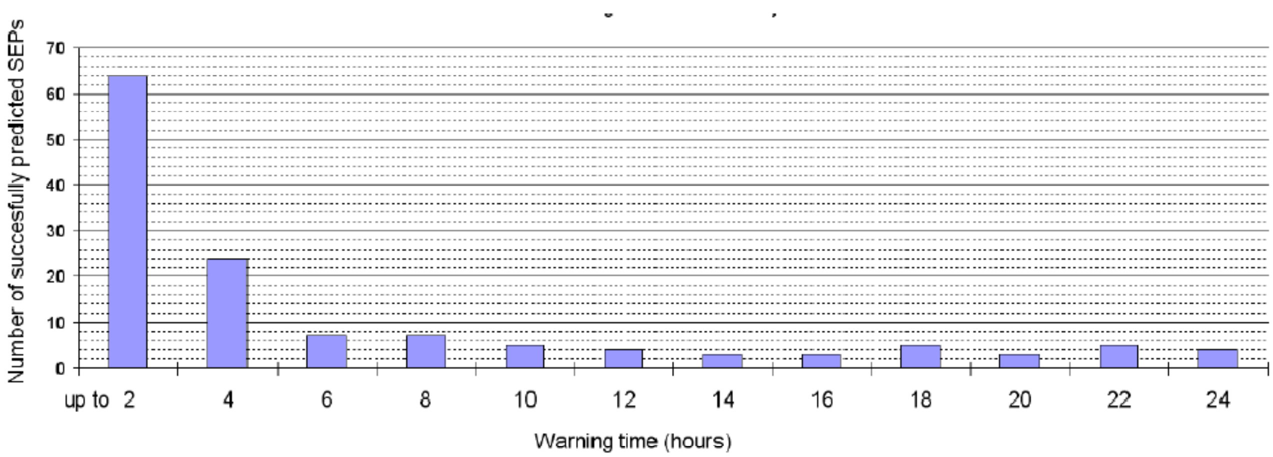

Figure 8. Distribution of warning times of the official UMASEP system using data from solar cycles 22 and 23.

For experiment 1, we used a list of 7 POD values[4]. To determine our best hypothesis regarding POD, we tested whether the most favorable hypothesis (i.e., POD>99 %) was not rejected. If it was rejected (as it obviously was), the percentage was reduced iteratively until the test was not rejected, arriving at the resulting hypothesis. The same procedure was applied to the FAR; if the most favorable hypothesis was rejected (i.e., FAR<1 %), the percentage was iteratively augmented until the test was not rejected. The results of experiment 1 were that the hypotheses POD>67.9% and FAR<31.3% are statistically significant. The results of experiment 2, with the official UMASEP using data from solar cycles 22 and 23, were that the hypotheses POD>75.0% and FAR<38.1% are statistically significant. For both experiments the significance level was 5%; that is, the confidence of the hypotheses was 95%. Current validations of SEP forecasters do not include statistical significance tests, so we only report our results and do not make any comparison on this topic; however, these results could be useful in the future for comparing our forecaster with others that use significance tests.

Figure 7 shows the overall performance of the official UMASEP system in terms of the number of correct SEP forecasts, missed SEPs and false alarms. Figure 8 shows the distribution of warning times for solar cycles 22 and 23.

[4] Section A.1 of the Auxiliary Material presents the lists of POD and FAR values that support the statistical significance tests mentioned in this section.

Section A1 of the auxiliary material includes two additional charts with the average warning time and the intensity error of UMASEP's forecasts from the data of solar cycles 22 and 23.

5. Comparisons with other forecasters

The comparison criteria for every case are expressed in terms of the reported POD and FAR for the current version of UMASEP and the POD and FAR of the aforementioned systems. All designers have reported either the POD and FAR or the necessary counters or timing data to calculate POD and FAR.

We compare systems in terms of the *all-type* POD_E, which is the probability of detecting *all types* (prompt and delayed) of SEP events in the energy range E; and, *all-type* FAR_E, which is the false alarm rate while predicting *all types* of (prompt and delayed) SEP events, taking into account the proton fluxes within the energy band E. If E is not indicated, we assume E > 10 MeV. For example, *all-type* $POD_{30\text{-}50MeV}$ will be the probability of detecting prompt and delayed 30-50 MeV SEP events.

This section is organized as follows: Section 5.1 compares our forecaster with four operative SEP forecasters [*Balch,* 2008; *Kahler et al.,* 2007; *Laurenza et al.*, 2009; *Posner,* 2007]. We also included a comparison preliminary version [*Núñez*, 2009, *Núñez and Núñez,* 2009] of the presented forecaster that we will call USF0.5; Section 5.2 presents some conclusions and possible explanations regarding the comparison results. The following paragraphs summarize the methods examined:

- The method proposed by Balch [2008] assumes that there is a relationship between the intensity of solar flare emissions and SEP event occurrence. Balch's program, called PROTONS, is based on the soft X-ray peak flux and time-integrated flux, the occurrence or non-occurrence of type II (associated with CME-driven shocks) and/or type IV radio bursts, and the Hα flare location.
- Kahler et al. [2007] developed a method, called PPS, for predicting solar energetic proton events by analyzing the solar flare peak, time-integrated X-ray fluxes, radio fluxes and times of onsets and maxima, and solar flare locations.
- Laurenza's approach [*Laurenza et al.*, 2009] is based on flare location, flare size, and evidence of particle acceleration/escape as parameterized by flare longitude, time-integrated soft X-ray intensity, and time-integrated intensity of type III radio emissions at 1 MHz, respectively. In this technique, warnings are issued 10 minutes after the maximum of >M2 soft X-ray flares.
- *Posner* [2007] developed an electron-based SEP prediction technique that exploits the shorter transit time of electrons relative to ions. This approach is based on the instrument COSTEP (SOHO), which provides data on relativistic electrons and <50 MeV protons (0.9 AU). This approach is specialized for forecasting SEPs in the range 30-50 MeV. Posner's SEP predictions will be available as long as the SOHO satellite is operative.
- The version USF0.5 [*Núñez,* 2009; *Núñez and Núñez*, 2009] was based on the same strategy as the well-connected forecasting model presented in this paper; however, USF0.5 has a slightly different correlation

Table 4. Performance of UMASEP for specific types and sizes of events, and levels of solar activity.

		Category of SEP event[a]	SEP peak flux level (pfu)	Probability of Detection (POD)	Average warning time	RMS error of ($\log_{10}$) intensity 7 h after onset
a)	All SEPs	Severe or extreme	$\times 10^4$	89.89% (8/9)	3 h 59 min	0.961
		Strong	$\times 10^3$	71.43% (15/21)	4 h 22 min	0.586
		Moderate	$\times 10^2$	89.74% (35/39)	4 h 29 min	0.350
		Minor	$\times 10$	78.35% (76/97)	6 h 8 min	0.342
		All		80.72% (134/166)	5 h 10 min	0.409

	Solar activity	Probability of Detection (POD)	False Alarm Rate (FAR)	Average warning time	RMS error of ($\log_{10}$) intensity 7 h after onset
b)	NSS > 150	70.00% (21/30)	25.93% (7/27)	6 h 35 min	0.406
	NSS = (100..150]	80.26% (61/76)	35.05% (34/97)	5 h 25 min	0.483
	NSS = (50..100]	80.50% (33/40)	40% (22/55)	4 h 22 min	0.314
	NSS≤ 50	95.00% (19/20)	25.00% (6/24)	5 h 11 min	0.334
	All activity levels	80.72% (134/166)	33.99% (69/203)	5 h 10 min	0.409

[a] During solar cycles 22 and 23, there were no "extreme" SEPs, which is an additional category ($\times 10^5$).

Table 5. Comparison tables among empirical SEP forecasters of >10 MeV SEP events

	SWPC SEP events[c] (1986-2004)	
	All-type POD[a]	All-type FAR[b]
PROTONS	57%	55%
UMASEP	79.61% (125/157)	34.89% (67/192)

a)

	SWPC SEP events[d] (1997-2001)	
	All-type POD[a]	All-type FAR[b]
PPS	40%, 56%[d] $\left(\frac{18}{45}\right), \left(\frac{57}{102}\right)$	50% (18/36)
UMASEP	77.77%[d] (35/45)	30%[d] (18/60)

b)

	SWPC SEP events + additional events[e] (1995-2005)	
	All-type POD[a]	All-type FAR[b]
Laurenza et al's approach	62.67% (47/75)	41.97% (34/81)
UMASEP	82.67% (62/75)	29.72%. (33/111)

c)

[a] Percentage of all types of SEP events correctly predicted. POD = number of correctly predicted events/number of all types of SEP events. The higher the POD, the better. If no energy range *E* is indicated in POD_E, it is assumed that the POD was calculated while predicting SEP events with energies E >10 MeV.
[b] Percentage of forecasts for which no SEP event occurred. FAR = number of forecasts for which no SEP event occurred/number of forecasts. The lower the FAR, the better.
[c] The reported POD and FAR of PROTONS correspond to the performances of the automatic empirical system at SWPC. The final yes/no predictions of SWPC NOAA are made by a human expert.
[d] Although the PPS system runs with >M5 flares, *Kahler et al* [2007] reported the *all-type* Hits and *all-type* Misses counters (for all sizes of flares), which facilitated the estimation of the *all-type* POD and *all-type* FAR. PPS also reported the *accuracy*, which is similar to the POD. To present the *all-type* POD of PPS, the calculated *all-type* POD was 40% (18/45) and the reported accuracy 56% (57/102) are presented.
[e] The validation set was composed of 66 SWPC SEPs and *additional* SEPs. From the seven *additional* SEPs, seven incremented the *Miss* events of UMASEP by 7.

measure and have neither the Analysis and Inference Module nor the poorly–connected SEP forecasting model.

To compare UMASEP with the empirical forecasters that used the SWPC SEP event list as their major reference [*Balch,* 2008; *Kahler et al*., 2007; *Laurenza et al.,* 2009; *Núñez and Núñez,*. 2009], we calculated new PODs and FARs for UMASEP with the dataset conditions that the compared systems used to calculate their PODs and FARs, making adjustments when needed as follows. For the cases we knew which SWPC SEP events were taken into account by the compared forecaster, we took them as reference for calculating our POD and FAR. For the rest of the validation situations we assumed our worst forecasting case scenario, that is: for the cases where *n* non-official SWPC SEP events were added to the list of events to be validated, we incremented our *miss* counter by *n,* as assuming that we were not able to predicted them; for the cases where *m* SWPC SEP were not considered by the other forecaster and their identification was not given, we reduced our *hit* counter by *m* for the analyzed period, as assuming that we predicted them but, now, we had to ignore them all. Table 2 presents the SWPC SEP events with the forecasting result of UMASEP for each of the events of the solar cycles 22 and 23. In the case of Posner's approach, which successfully predicts 30-50 MeV SEP events, we wanted to know its skill for forecasting >10 MeV SEP events, so users could know more about this system by forecasting well-known and hazardous SWPC SEP events. We estimated the *all-type* $POD_{>10MeV}$ and *all-type* $FAR_{>10MeV}$ for 2003 by comparing the times of the forecasts, presented in [*Posner*, 2007], with the official start times of the SWPC SEP events.

5.1 Comparison of UMASEP with automatic forecasters of >10 MeV SEP events

Regarding the PROTONS program [*Balch,* 2008], the reported POD for predicting all types (prompt and delayed) of events, abbreviated as *all-type POD* throughout this section, was 57%. The FAR of PROTONS during the period 1986 to 2004 while predicting all types of events, abbreviated as *all-type* FAR throughout this section, was 55%. According to Table 2, UMASEP had an *all-type* POD of 79.61% (125/157) and an *all-type* FAR of 34.89% (67/192) for the same period when the same official SWPC SEP list was used. These results are summarized in Table 5a. If the POD and/or FAR of both systems were similar, no conclusion could be made. However, the results are not similar, we therefore conclude that UMASEP has a better POD and FAR than the PROTONS program.

It is important to mention that the PROTONS program is used as a decision aid for human SEP forecasting experts, however the final yes/no SEP prediction is made by the human experts. For the interval from 1995 to 2005, the SWPC forecasting infrastructure had an *all-type* POD of 87.64% (78/89) with an *all-type* FAR of 17.89% (17/95) [*Balch,* 2008]. According to Table 2, during that period, UMASEP had an *all-type* POD of 84.27% (75/89) with an *all-type* FAR of 24.77% (28/113). Therefore, the NOAA/SWPC forecasting performance (including human experts and the PROTONS program) yields better results than the automatic UMASEP system.

Kahler et al. [2007] validated the PPS system using data from 1997 to 2001. There are 50 SEP events in the SWPC SEP list for the period 1997-2001. PPS added two unlisted SEP events that occurred during high (>10 pfu) GOES intensities and subtracted seven SEP events. For the remaining 45 events, they successfully predicted 18 (namely A). They also reported 18 false alarms (namely B), 27 missed events (namely C) and 39 *null* events that successfully predicted (namely D). The performance measure that they reported is known as *accuracy*, which is (A+D) / (A+B+C+D). The reported accuracy of PPS is 56% ((18+39)/(18+18+27+39)). The *all-type* POD and *all-type* FAR may be directly calculated from the A, B and C counters. Using the same counters, the *all-type* POD of PPS was A/(A+C), that is, 40% (18/(18+27)), and the *all-type* FAR of PPS was B/(A+B), that is, 50% (18/(18+18)). Because PPS added two SEP events and subtracted seven events, we assume our worst forecasting case scenario: UMASEP fails to predict the two new SEP events that PPS added to the analysis (our miss counter is incremented by two), and does not predict the seven discarded SEP events (our *hit* counter is decremented by seven). Our original *all-type* POD was 42/50 (without making adjustments) during the same period (1997-2001), according to Table 2, so the new POD is (42-7)/(50+2-7), that is 77.77% (35/45). UMASEP had an *all-type* FAR of 30% (18/60). These results are summarized in Table 5b. One of the reasons for the difference in performance between PPS and our forecaster (77.77% vs. 40%, and 30% vs. 50%) is that PPS has a good performance with large flares (>M5 flare), while UMASEP has the best performance with >C7 flares, which allows the forecasting of more events. If the POD and/or FAR of both systems were similar, no conclusion could be made. However, the results are not similar, we therefore conclude that UMASEP has a better POD and FAR than the PPS program.

Laurenza et al. [2009] validated their system with 75 SEP events from 1995 to 2005. From the original 93 SWPC SEP events, they considered 68 SWPC SEP events. They excluded events with certain conditions (e.g., those for which the responsible eruption was located on the backside of the Sun and the associated SXR flare was <M2, and those for which the SXR or radio data were unavailable). They also added 7 events, mostly strong proton enhancements occurring with strong flares in the middle of a declining (but still valid >10 pfu) event, (e.g., April 12, 2001, October 29, 2003). *Laurenza et al*. [2009] reported an *all-type* POD of 62.67% (47/75) and an *all-type* FAR of 41.97% (34/81). To compare the systems, we calculated the POD of our predictor by taking into account only Laurenza et al.'s SEP events. We counted as missed events the added 7 Laurenza et al's events. Our original POD on the same 68 SWPC events, was 62/68 (according to Table 2 and Figure 7); however, we added 7 misses, obtaining 62/(68+7) as the new *all-type* POD of UMASEP, that is, 82.67% (62/75). The all-type FAR is not affected with the added Laurenza et al's events, so according with the hit and false-alarm counters in Figure 7, the *all-type* FAR of UMASEP for the period from 1995 to 2005 was 29.72% (33/111), which is not affected by the added misses. These results are summarized in Table 5c. One of the reasons for the difference in performance between the forecasters developed by us and Laurenza et al. is the condition of working only with >M2 flares. UMASEP may issue predictions with lower-size flares (>C7), which allows the forecasting of more events. If the POD and/or FAR of both systems were similar, no conclusion could be made. However, the results are not similar, we therefore conclude that UMASEP has a better all-type POD and all-type FAR than Laurenza et al's program.

Regarding the previous version of UMASEP, USF0.5 [*Núñez,* 2009; *Núñez and Núñez*, 2009], for solar cycles 22 and 23, our current version of the forecaster achieved a *prompt SEP POD* (the probability of detecting prompt SEPs) of 89.16% (74/83) compared to 84 % (70/83) in the current version of UMASEP. For the same period, the *prompt FAR* of USF0.5 was 46.15% (60/130) as compared with 33.99% (69/203) of the current version of UMASEP. Because prompt SEPs constitute half of the all-type SEPs, the *all-type* POD of USF0.5 is 44.57% (74/166) and the *all-type* FAR is 46.15% (60/130). Taking into account the above results, we conclude that the current version of UMASEP outperforms the previous version in terms of *all-type* POD and *all-type* FAR. Because these results belong to a different version of UMASEP, no summary is presented in Table 5.

5.2. Comparison of UMASEP with an automatic forecaster that predict 30-50 MeV SEP events

SWPC uses the threshold J (E>10 MeV) = 10 pfu for identifying SEP events. SWPC SEP threshold is used with the same purpose globally. Posner's approach [*Posner*, 2007] is specialized in predicting 30-50 MeV SEP events. We want to know its skill for forecasting

>10 MeV SEP events, so users could know more about this system by forecasting well-known and hazardous SWPC SEP events. For example, during 2003, October 26th, users needed to be warned before the occurrence of the first Halloween event by any forecaster, independently of the forecaster's particularities (i.e. target proton energy band, conditions for running, etc), because these predictions were important to take important decisions. We want to estimate the *all-type* $POD_{>10\ MeV}$ and *all-type* $FAR_{>10\ MeV}$ of Posner's approach by comparing the times of its forecasts with the official start times of the SWPC SEP events. Table 6 presents data in chronological order. Each row of the table could show either a time associated to a SWPC SEP event (start or end times) or the time of a forecast issued by either the Posner's approach or UMASEP. Depending on the type of information, the columns are filled in: if the event is a SWPC SEP event, the second column ("SWPC SEP Event Information") is filled in; if the row is a forecast issued by the Posner's approach, the third and fourth columns are filled with the forecast details and results; if the forecast is issued by UMASEP, the fifth o sixth columns are filled in with the forecast details and results. Finally, the last column ("Earliest forecaster") summarizes which forecaster better anticipated the corresponding >10 MeV SEP event. A blank indicates that the datum associated with the column (SEP event or forecast) did not occur at the corresponding time.

For all SEP events of Table 6 we consider that a forecast is a $Hit_{>10\ MeV}$ when the time of the forecast is lower than the start time of the SWPC SEP event. If the time of the forecast is equal or larger than the Start Time of the SEP event, a $Miss_{>10MeV}$ occurs. When there is a $Hit_{>10MeV}$, the warning time is also shown. According to Table 6, the all-type $POD_{>10MeV}$ of Posner's approach was 37.50% (3/8) against 75.00% % (6/8) of UMASEP. The all-type $FAR_{>10MeV}$ of Posner's approach was 40.00% (2/5) against 40.00% (4/10) of UMASEP. UMA SEP better anticipated five SEP events, while Posner better anticipated two events. For example, the integral proton flux (E >10 MeV) of the very energetic prompt SEP of 2003, October 28th, surpassed the 10 pfu threshold earlier than the onset of the 30-50 MeV proton event. Given the forecast times (UT) of Table 6, and the timing data shown in Figure 2b, the SEP event 2003 October 28 was predicted by UMASEP (with the GOES P7 channel) 47 minutes before Posner's approach (with the SOHO/COSTEP instrument). The results show that >10 MeV SEPs are preceded, not only by relativistic electrons but also by low fluxes of very energetic protons. The challenge is how to process such small fluxes. This paper proposes a satisfactory method in which to do so.

In summary, UMASEP obtained a better *all-type* $POD_{>10MeV}$ (75% against 37.5% of Posner's approach) and similar *all-type* $FAR_{>10MeV}$ (40%) for predicting SEP events of 2003. UMASEP anticipated more events, 5 events compared to 2 of Posner's approach. By taking into account these comparisons we conclude that the UMASEP had a better overall performance than Posner's approach for predicting >10 MeV SEP events.

To find explanations regarding the presented results, we noted that from 65 SEP events (E > 10 MeV) occurred during 1998-2002 shown in Table 1a of *Posner* [2007], 17 were not preceded by 0.3-1.2 MeV relativistic electrons. That is, 26.15% (17/65) of the SWPC SEP events would have been missed by a forecasting strategy that uses 0.3-1.2 MeV electrons. In addition to that, during the same period 9% (6/65) of the SWPC SEP events of this period, were detected before the 30-50 MeV SEP onset, which complicates the forecasting results of this strategy for predicting >10MeV events. These 6 SWPC SEP events (e.g., 2001, October 24 and 2001, April 15) were identified before the onset of 30-50 MeV SEP events by comparing the official start time of > 10 MeV SEPs (http://www.swpc.noaa.gov/ftpdir/indices/SPE.txt) with the 30-50 MeV proton onset (begin) time of Table 2a of Table 2a in Posner[2007].

Posner's approach is more closely related to our system than PROTONS, PPS and Laurenza's system because it has a similar strategy: analyzing particle fluxes in the near-Earth environment to predict hazardous proton events. A difference in ours is that we analyze differential proton and soft X-ray fluxes. A disadvantage of Posner's approach compared with the other forecasters is the low availability (60% or even less) of the communication link with the data source (SOHO/COSTEP). This availability reduces the probability of detection of SEP events of Posner's approach in real-time.

5.3. Additional comments

Because damage to equipment or health problems for astronauts might occur due to SEP events, space weather users want to be warned against these events, regardless of the limitations of the forecasting infrastructure (space instruments and forecaster programs).

Table 6. Comparison Between *Posner*'s [2007] Approach and UMASEP for Predicting >10 MeV SEP Events for the Year 2003[a]

Time[b]	SWPC SEP Event Information[c]	Posner's Approach Outputs and Results: Forecast Occurrence[d]	Posner's Approach Outputs and Results: Forecast Result[e]	UMASEP Outputs and Results: Forecast Occurrence	UMASEP Outputs and Results: Forecast Result[e]	Earliest Forecaster of the >10 MeV Event[f]
3/17 19:15		Forecast issued (end time: 3/17 19:15)	$\text{FalseAlarm}_{>10MeV}$			
5/28 11:55				Forecast issued (end time: 5/29 11:55)	Successful	
5/28 23:35	SEP start time		$Missed_{>10MeV}$		$Hit_{>10MeV}$ (11 h 40 min)	UMASEP
5/29 12.22		Forecast issued (end time: 5/29 18:12)	Ignored[g]			
5/30 02:40	SEP end time					
5/31 03:05				Forecast issued (end time: 5/31 11:55	Successful	
5/31 03:18		Forecast issued (end time: 5/31 08:58)	Successful			
5/31 04:40	SEP start time		$Hit_{>10MeV}$ (1 h 22 min)		$Hit_{>10MeV}$ (1 h 35 min)	UMASEP
5/31 16:25	SEP end time					
6/1 16:05				Forecast issued (end time: 6/01 21:10)	$\text{FalseAlarm}_{>10MeV}$	
6/2 17:55				Forecast issued (end time: 6/02 20:15)	$\text{FalseAlarm}_{>10MeV}$	
6/18 16:05				Forecast issued (end time: 6/19 16:05)	Successful	
6/18 20:50	SEP start time		$Missed_{>10MeV}$		$Hit_{>10MeV}$ (4 h 45 min)	UMASEP
6/19 13:00	SEP end time					
10/23 13:45				Forecast issued (end time: 10/24 13:45	$\text{FalseAlarm}_{>10MeV}$	
10/26 17:54		Forecast issued (end time: 10/26 18:31)	Successful			
10/26 18:00				Forecast issued (end time: 10/26 23:10)	Successful	
10/26 18:25	SEP start time		$Hit_{>10MeV}$ (31 min)		$Hit_{>10MeV}$ (25 min)	Posner's approach
10/27 20:40	SEP end time					
10/28 11:40				Forecast issued (end time: 10/28 23:20)	Successful	
10/28 12:15	SEP start time		$Missed_{>10MeV}$		$Hit_{>10MeV}$ (35 min)	UMASEP
10/28 12:27		Forecast issued	Ignored[g]			
11/01 16:00	SEP end time					
10/31 06:26		Forecast issued (end time: 10/31 13:06)	$\text{FalseAlarm}_{>10MeV}$			
11/2 11:05	SEP start time		$Miss_{>10MeV}$		$Miss_{>10MeV}$	None
11/2 17:26		Forecast issued (end time: 11/02 19:19)	Ignored[g]			
11/4 21:00	SEP end time					
11/4 08:32		Forecast issued (end time: 11/05 03:05)	Successful			
11/4 22:25	SEP start time		$Hit_{>10MeV}$ (1 h 20 min)		$Miss_{>10MeV}$	Posner's approach
11/7 21:55	SEP end time					
11/20 10:30				Forecast issued (end time: 11/21 12:42)	$\text{FalseAlarm}_{>10MeV}$	
11/21 22:05				Forecast issued (end time: 11/22 22:54)	Successful	
11/21 23:55	SEP start time		$Missed_{>10MeV}$		$Hit_{>10MeV}$ (1 h 50 min)	UMASEP
11/21 23:58		Forecast issued (end time: 11/22 06:00)	Ignored[g]			
11/22 22:20	SEP end time					

[a]The SWPC SEP events for the year 2003 are presented in italic font.
[b]Date format is month/day. Time is in UT.
[c]Start times (*ST*) of SEP events are presented according to the SWPC SEP event list (http://www.swpc.noaa.gov/ftpdir/indices/SPE.txt).
[d]Begin and end time of 30–50 MeV SEP forecasts according to *Posner* [2007]. Date format is month/day. Time is in UT.
[e]The forecast result of forecasting SEP events with the SWPC threshold J (E > 10 MeV) ≥ 10 pfu. When there is a $Hit_{>10\ MeV}$, the anticipation time (hours) is also presented.
[f]The forecaster with the highest warning time (shown with $Hits_{>10MeV}$) correspond to the forecaster with better anticipation for the indicated >10 MeV SEP event.
[g]The forecast is issued while a SEP is occurring. This delayed forecast of Posner's approach is not a false alarm because it does not unnecessarily warn users (the event is already occurring) or a successful forecast because it was late. These warnings are ignored, so they do not affect the counters for calculating the POD and FAR of Posner's approach.

UMASEP can issue predictions for more SEPs than other forecasters [*Laurenza et al.*, 2009; *Balch,* 2008; *Kahler et al.,* 2007], particularly for those events associated to flares with X-ray flux intensities from C7 to M2.

The methods that make predictions mainly based on solar data (PROTONS, PPS and Laurenza *et al*'s approach) have better warning time for delayed SEP events than the methods that process particle fluxes in the near Earth (Posner's approach and UMASEP). For example, *Laurenza et al*'s program issues forecasts 10 minutes after the maximum of >M2 soft X-ray flares. According to our calculations only for the delayed SEP events of their list of events, the average warning time of Laurenza et al.'s approach for these SEP events is 18 h 23 min. The warning time of UMASEP for the same set of delayed events was 7 h 12 min. We therefore conclude that Laurenza *et al.*'s approach has better average warning time than UMASEP when delayed SEP events are successfully predicted. Because their similar strategy, we infer that the methods that predict SEPs from solar data (PROTONS, PPS and Laurenza et al's approach) have better anticipation than UMASEP and Posner's approach for poorly-connected SEPs, when they are correctly predicted.

Posner's approach has achieved a high performance (high $POD_{30\text{-}50\ MeV}$ and very low $FAR_{30\text{-}50\ MeV}$) for predicting 30-50 MeV SEP events. Although it was not tuned with >50 MeV proton data, Posner's approach successfully forecasted 3 out 8 SWPC SEP events, and was near to anticipate 2 more events, showing that a future and specific >10 MeV-based tuning would yield better results for predicting >10 MeV SEP events, in addition to the prediction of its current target events. On the other hand, its low false alarm rate, similar to that of UMASEP (probably because of a similar strategy of processing particle fluxes near Earth), is an advantage that would easily allow Posner's approach to issue more warnings and therefore enhance the POD of >10 MeV events.

Predicting SEP events only by processing solar data is a difficult task. Several phenomena may occur in different parts of the Sun at similar times, causing the interplanetary medium to be dynamic, so it is not easy to predict whether solar protons can escape to the interplanetary medium and reach Earth, surpassing the official SWPC SEP threshold at 1 AU. On the other hand, approximately one fourth of >10 MeV SEP events are not preceded by 0.3-1.2 MeV relativistic electrons. Unlike other methods, UMASEP processes low differential proton fluxes to encounter symptoms that more protons could reach Earth. This strategy allows our forecaster to filter out potential false alarms more easily. The higher $POD_{>10MeV}$ and lower $FAR_{>10MeV}$ of UMASEP compared with the rest of the empirical forecasters is mainly due to its capacity for processing low differential proton fluxes.

6. Future lines of research and conclusions

The UMASEP system makes real-time predictions of the time interval within which the integral proton flux is expected to meet or surpass the SWPC SEP threshold and the intensity of the first hours of SEP events. Taking into account data from solar cycles 22 and 23, the UMASEP had an *all-type* $POD_{>10\ MeV}$ of 80.72%, an *all-type* $FAR_{>10\ MeV}$ of 33.99% and an average warning time of 5 h 10 min (1 h 5 min for well-connected events and 8 h 28 min for poorly-connected events, with a maximum of 24 hours for the case of very gradual SEP events). The warning time is the temporal distance between the time when the prediction is issued and the time when the integral proton flux meets or surpasses the SWPC SEP threshold.

The well-connected SEP forecasting model tries to estimate the magnetic connectivity by recognizing a cause-effect pattern between the associated CME/flare zone and the Earth (in terms of soft X-rays), and the near-Earth environment (in terms of differential proton fluxes). If a magnetic connection is present and persists, and if it is followed by a C7 flare or higher, this model predicts a well-connected SEP event. The second model analyzes the differential proton fluxes; it tries to detect whether the differential proton flux behavior is similar to the beginning phase of previous historically poorly-connected SEP events, and if so it deduces similar consequences. If the predicted event is well-connected, the system gives information about the associated flare and the active region and/or connected heliolongitude.

Among all SEP forecasting approaches (automatic systems and human experts assisted by automatic systems), the NOAA Space Weather Prediction Center (composed of the PROTONS program and human SEP forecasting experts) performs better that the automatic UMASEP system in terms of POD (SWPC 87.64%, UMASEP 84.27%) and FAR (SWPC 17.89%, UMASEP 24.77%) in the forecasts of all types of SEPs for the period 1995-2005, as reported by *Balch* [2008].

The UMASEP system was compared with several well-known automatic forecasters [*Balch*, 2008; *Kahler et al.*, 2007; *Laurenza et al*., 2009; *Posner*, 2007; *Núñez and Núñez*, 2009]. We calculated new PODs and FARs for UMASEP with the dataset conditions that the compared systems used to calculate their PODs and FARs, making adjustments when needed as follows. For the cases we knew which SWPC SEP events were taken into account by the compared forecaster, we took them as reference for calculating our POD and FAR. For the rest of validation situations we made the following assumptions: for the cases where *n* non-official SWPC SEP events were added to the list of events to be validated, we incremented our *miss* counter by *n,* as assuming that we were not able to predicted them; for the cases where *m* SWPC SEP were not considered by the other forecaster and their identification was not given, we reduced our *hit* counter by *m* for the analyzed period, as assuming that we predicted them but, now, we had to ignore them all. We also wanted to know the skill of Posner's approach for forecasting >10 MeV SEP events, so we estimated its *all-type* $\mathrm{POD}_{>10\mathrm{MeV}}$ and *all-type* $\mathrm{FAR}_{>10\mathrm{MeV}}$ for 2003 by comparing the times of the forecasts, presented in [*Posner*, 2007], with the official start times of the SWPC SEP events. Based on these comparison criteria, we concluded that UMASEP had the best probability of detecting all type (prompt and delayed) of SEP events, also called *all-type* $\mathrm{POD}_{>10\mathrm{MeV}}$, and the best *all-type* $\mathrm{FAR}_{>10\mathrm{MeV}}$, of all the analyzed empirical forecasters that predicts events according to the SWPC SEP threshold. The second best overall performance of the analyzed empirical forecasters was Laurenza et al's approach. For the period 1995-2005, Laurenza et al.'s approach had an *all-type* $\mathrm{POD}_{>10\mathrm{MeV}}$ of 62.67% compared to 82.67% of UMASEP, and an *all-type* $\mathrm{FAR}_{>10\mathrm{MeV}}$ of 41.97% compared to 29.72% of UMASEP for the period 1995-2005 (see Table 5c).

An interesting finding arisen from the detailed temporal comparison between the Posner's approach and UMASEP. The second Halloween event, the prompt and most energetic event of 2003, was forecasted by UMASEP (specialized in >10 MeV events) 47 minutes earlier than Posner's approach (specialized in 30-50 MeV events). We compared the official start times of SWPC SEP events for the period 1998-2002, with the corresponding onset times of 30-50 MeV proton events [*Posner*, 2007], and we realized that in 9% of SWPC SEP events, their start time (time of meeting or surpassing the SWPC SEP threshold) occurred before the onset time (time of the first rises) of the corresponding 30-50 MeV SEP events. For the aforementioned event of 2003, October 28, *low* fluxes of very energetic 165-500 MeV solar protons (corresponding to the GOES P7 channel which UMASEP analyzed) arrived earlier than *large* fluxes of 0.3-1.2 MeV relativistic electrons (detected by SOHO-COSTEP whose data is analyzed by Posner's approach). We could say, in general that prompt hazardous fluxes of solar protons are preceded not only by relativistic electrons but also by low fluxes of solar protons in specific high-energy bands; the challenge is how to process such low fluxes. This paper empirically proves that correlating them with X-ray fluxes may be useful to deduce magnetic connections and therefore to predict prompt hazardous fluxes of solar protons.

An advantage of the methods that make predictions mainly based on solar data (PROTONS, PPS and Laurenza *et al*'s approach) over the methods that process particle fluxes in the near Earth (Posner's approach and UMASEP) is the average warning time when they successfully predict delayed SEPs, because they are able to predict these events minutes after the major flare has occurred. However, given the POD and FAR of all analyzed methods, we say that predicting SEP events by processing only solar data is a difficult task. Several phenomena may occur in different parts of the Sun at similar times, causing the interplanetary medium to be dynamic, so it is not easy to predict whether solar protons can escape to the interplanetary medium and reach Earth, surpassing the official SWPC SEP threshold at 1 AU. On the other hand, approximately one fourth of >10 MeV SEP events are not preceded by 0.3-1.2 MeV relativistic electrons. Unlike other methods, UMASEP processes low differential proton fluxes to encounter symptoms that more protons could reach Earth. This strategy allows UMASEP to have a low FAR, and therefore reduce the flare size limitations (i.e. >C7 flares for UMASEP instead of M flares for other forecasters) which enhances the POD. By analyzing 10-500 MeV differential proton fluxes at 1 AU (including low fluxes), UMASEP may issue well-connected SEP forecasts when no enhancement in the integral proton flux (E >10 MeV) is observed (see Figures 2 and 4), and some times, when no strong 30-50 MeV fluxes have been detected. In the case of poorly-connected events, stronger proton flux rises need to be analyzed, including integral proton flux enhancements.

Differential proton fluxes had not been used by empirical SEP forecasters. This paper shows that the real-time analysis and/or correlation of the differential proton fluxes (in particular low fluxes), with solar phenomena not only carry very important information about the Sun-Earth link status, but also that they can be used for predicting SEP events. The satisfactory results of the well-connected SEP forecasting model and its necessary condition of the existence of magnetic connections between the observer and the X-ray process might imply that there also exist intermediate magnetic connections from the flare site to the CME-driven shock during the first hours of the proton enhancement, so fresh suprathermal protons reach the shock and finally are reaccelerated by it. This possibility is considered physically possible, but not yet proven [Tylka et al. 2005]. Our results might also be explained by taking into account other solar processes (e.g., Chen and Kunkel [2010] found evidence that there is a physical relationship between flare energy release and poloidal magnetic flux injection in CMEs). However, no conclusion can be drawn about these flare/CME processes, and more studies need to be performed by the scientific community. On the other hand, our model has the limitation of not analyzing the shock. For this reason, future studies of including real-time analyses of signatures of acceleration at the CME-driven shock (e.g., type II radio emissions [*Holman and Pesses*, 1983]), as other methods do, seem to be an interesting field of research for refining our SEP forecasting models by inferring better conclusions about the situations.

UMASEP is the only system that automatically identifies the associated flare and active region. It is also the only system that was evaluated for two solar cycles and that may be evaluated by any user using the public service (http://spaceweather.uma.es/uma_sep_tool.htm). A user may submit a file with X-ray and differential proton flux data as input to UMASEP and observe the corresponding outputs to evaluate specific situations or to complement his/her studies with the forecaster's hypotheses about magnetic connections that are not necessarily associated with solar proton events.

UMASEP's approach could be used on-board a spacecraft with detectors of soft X-ray and differential proton fluxes from 10 to 500 MeV. Radiation risk can be significantly mitigated by on-site forecasting of SEP events for voyages to the Moon, the asteroids or Mars. Magnetic connections associated with Earth might drastically differ from the magnetic connections associated with an interplanetary spacecraft; therefore, if a hazardous SEP event is well-connected, it could be so close that an on-site forecast could be the only warning message for autonomous control modules, or astronauts, to take the proper preventive actions. The high average warning time, the high probability of detection and the low false alarm rate make this forecaster useful for space users to prevent equipment and humans from hazardous solar proton events.

Acknowledgements

This work was partially funded by project TIC07-2861 of the Junta de Andalucía, Spain. I would like to express my sincere gratitude to Karl-Ludwig Klein, Kamen Kozarev and Raúl Fidalgo for their careful revisions and comments. I also want to thank Michael Hesse for his support and encouragement and David Núñez for his contribution during the system evaluation phase and his insightful revision of the paper.

References

Aran, A., B. Sanahuja and D. Lario, (2006), SOLPENCO: A SOLar Particle ENgineering COde, Advances in Space Research, Vol. 37, Issue 6, pp 1240-1246, doi:10.1016/j.asr.2005.09.019.

Arge, C. N. and V. J. Pizzo (2000), Improvement in the prediction of solar wind conditions using near-real time solar magnetic field updates, Journal of Geophysical Research, 105, 10465–10480.

Balch, C. C. (2008), Updated verification of the Space Weather Prediction Center's solar energetic particle prediction model, *Space Weather*, 6, S01001, doi:10.1029/2007SW000337.

Beck, P., M. Latocha, S. Rollet, and G. Stehno (2005), TEPC reference measurements at aircraft altitudes during a solar storm, *Adv. Space Res.*, 16(9), 1627-1633, doi:l0.10l6/j.asr.2005.05.035.

Belov A., Properties of solar X-ray flares and proton event forecasting, *Advances in Space Research*, Volume 43, Issue 4, p. 467-473. 2009.

Boberg,, F., P. Wintoft and H. Lundstedt (2000), Real time Kp predictions from solar wind data using neural networks, in Physics and Chemistry of Earth, Vol. 25, No 4.

Breiman, L. and J. Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984) Classification and Regression Trees, Wadsworth & Brooks, 1984.

Cane, H. V., T. T. von Rosenvinge, C. M. Cohen, and R. A. Mewaldt. (2003), Two components in major solar particle events, *Geophysical Research Letters*, 30(12), 8017, doi:10.1029/2002GL016580.

Cane, H. V., R. A. Mewaldt, C. M. S. Cohen, and T. T. von Rosenvinge (2006), Role of flares and shocks in determining solar energetic particle abundances, J. Geophys. Res., 111, A06S90, doi:10.1029/2005JA011071.

Chen J. and V. Kunkel (2010), Temporal and physical connection between coronal mass ejections and flares, *Astrophys. J.*, 717, 1105-1122, doi: 10.1088/0004-637X/717/2/1105.

Davis, J., and M. Goadrich, (2006), The relationship between Precision-Recall and ROC curves, *Proceedings of the 23rd international Conference on Machine Learning* (Pittsburgh, June 25 - 29, 2006). ICML '06, vol. 148. ACM, New York, 233-240. DOI= http://doi.acm.org/10.1145/1143844.1143874

Deng, Y., J. Xing and Zhong Y. (2005), Ensemble SVR for prediction of time series, Proceedings of International Conference on Machine Learning and Cybernetics, Vol. 6, pp 3528-3534.

Di Marco, F., A. Donati, M. Pantoquilho, A. Baumgartner, and D. Ponz (2008), SEISOP: Space Environment Information System to support Satellites Operations, Proceedings of the SpaceOps 2008 Conference.

Hansen, L. K. and P. Salamon (1990) Neural network ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(10):993–1001, 1990.

Holman, G. and M. Pesses, (2003), Solar type II radio emission and the shock drift acceleration of electrons, Astrophysical Journal, vol. 267, p. 837-84.

Kahler, S.W., E.W. Cliver, and A.G. Ling (2007), Validating the proton prediction system, *Journal of Atmospheric and Solar-Terrestrial Physics*, Volume 69, Issues 1-2, 2007, Pages 43-49.

Klein, K.-L, S. Krucker, G. Trottet, and S. Hoang (2005), Coronal phenomena at the release of solar energetic electron events, Astron. Astrophys., 431, 1047-1060, doitl0.1051/0004-6361:20041258.

Koskinen, H., E. Tanskanen, R. Pirjola, A. Pulkkinen, C. Dyer, D. Rodgers, P. Cannon, J. C. Mandeville, and D. Boscher (2001), *Space weather effects catalogue*, Finnish Meteorological Institute, Reports, 2001:2.

Kubo, Y., and M. Akioka (2004), Existence of thresholds in proton flares and application to solar energetic particle alerts, Space Weather, 2, SOI002, doi:10.1029/2003SW000022.

Kuncheva, L. and C. Whitaker (2003), Measures of diversity in classifier ensembles, *Machine Learning*, 51, pp. 181-207, 2003

Laurenza, M., E. W. Cliver, J. Hewitt, M. Storini, A. G. Ling, C. C. Balch, and M. L. Kaiser (2009), A technique for short-term warning of solar energetic particle events based on flare location, flare size, and evidence of particle escape, *Space Weather*, 7, S04008.

Linker, J. A., P. Riley, Z. Mikic, R. Lionello, V. Titov, J. Wijaya (2010), CORHEL MHD Modeling in Support of Solar Dynamics Observatory, paper presented at the American Astronomical Society (AAS) Meeting #216, Bulletin of the American Astronomical Society, Vol. 41, p. 876.

Lionello, R., J.A. Linker, and Z. Mikic (2003), Three-dimensional Magnetohydrodynamic of the Solar Corona and the Solar Wind with Improved Energy Transport, *Proceedings of the Tenth International Solar Wind Conference*, American Institute of Physics Conference Proceedings 679. Melville, New York: American Institute of Physics, p. 222.

Lucci, N. (2005), Space weather conditions and spacecraft anomalies in different orbits, *Space Weather*, 3, S01001, doi:10.1029/2003SVV000056.

Marque, C, A. Posner, and K.-L. Klein (2006), Solar energetic particles and radio-silent fast coronal mass ejections, Astrophys.]., 642, 1222-1235, doi:10.1086/501157.

Maddox, M., M. Hesse, M. Kuznetsova, M. Rastaetter, L. MacNeice, P. J. Jain, W. Garneau, D.H. Berrios, A. Pulkinnen, and D. Rowland (2008), The Integrated Space Weather Analysis System, paper presented at the American Geophysical Union Fall Meeting 2008 (Website: http://iswa.gsfc.nasa.gov/iswa/iSWA.html).

Miller, J., C. Zeitlin, F.A.Cucinotta, , L. Heilbronn, , D. Stephens, and J. W. Wilson (2003), Benchmark Studies of the Effectiveness of Structural and Internal Materials as Radiation Shielding for the International Space Station, Radiation Research 159:381–390.

Nuñez, M, R. Fidalgo, M. Baena, R.Morales, The Influence of Active Region Information in the Prediction of Solar Flares: an Empirical Model using Data Mining, Annales Geophysicae, 23, 9 (2005) 3129-3138.

Núñez, M, R. Fidalgo and R. Morales (2006), On forecasting the onset of Solar Proton Events, in Proceedings of the IAU Symposium 233 Solar Activity and its Magnetic Origins, pp 81-82, Cambridge University Press.

Núñez, M. (2009), Real-Time Forecasting of SEP Events with Warning Time of up to Seven-Hours and Low False Alarm Rate, presented at Space Weather Workshop 2009, Boulder, Colorado.

Núñez, M., and D. Núñez (2009), A Model for Predicting the Onset and Intensity of Well-Connected SEPs and its Evaluation for Cycles 22 and 23, paper presented at the European Space Weather Week, Belgium.

Odstrcil, D., V. J. Pizzo, J. A. Linker, P. Riley, R. Lionello, and Z. Mikic (2004), Initial Coupling of Coronal and Heliospheric Numerical Magnetohydrodynamic Codes, *J. Atmos. Solar-Terrestrial Phys*., 66, 1321.

Parker, E. N. (1958), Dynamics of the Interplanetary Gas and Magnetic Fields, *Astrophysical Journal*, vol. 128, p.664.

Posner, A. (2007), Up to 1-hour forecasting of radiation hazards from solar energetic ion events with relativistic electrons, *Space Weather*, 5, S05001, doi:10.1029/2006SW000268.

Qahwaji, R. and T. Colak (2007), "Automatic Short-Term Solar Flare Prediction Using Machine Learning And Sunspot Associations," Solar Physics,241,pp. 195-211.

Quinlan, J.R. (1992), Learning with continuous classes. Proceedings 5th Australian Joint Conference on Artificial Intelligence, World Scientific pp 343-348.

Rapp, D. (2006), Radiation Effects and Shielding Requirements in Human Missions to the Moon and Mars, *The Mars Journal,* 2, 46-71, 2006; doi:10.1555/mars.2006.0004.

Reames, D. V. (2004), Solar Energetic Particle Variations, *Adv. Space Res*. 34 (No. 2), 381.

Schwadron, N. A., L. Townsend, K. Kozarev, M. A. Dayeh, F. Cucinotta, M. Desai, M. Golightly, D. Hassler , R. Hatcher, M.Y. Kim, A. Posner, M. PourArsalan, H. E. Spence, and R. K. Squier, (2010), Earth-Moon-Mars Radiation Environment Module framework, *Space Weather*, Vol. 8, S00E02, doi: 10.1029/2009SW000523.

Sokolov, I.V., Roussev, I.I., Gombosi, T.I., Lee, M.A., K6ta, J., Forbes, T.G., Manchester, W.B., Sakai, J. . (2004), A new field line advection model for solar particle acceleration. Astrophysical Journal 616 (2), L171-L174.

Toth, G., I. V. Sokolov, T. I. Gombosi, D. R. Chesney, C. R. Clauer, D. L. De Zeeuw, K. C. Hansen, K. J. Kane, W. B. Manchester, R. C. Oehmke, K. G. Powell, A. R. Ridley, I. I. Roussev, Q. F. Stout, O. Volberg, R. A. Wolf, S. Sazykin, A. Chan, B. Yu, J. Kota (2005), Space Weather Modeling Framework: A New Tool for the Space Science Community, *Journal of Geophysical Research,* 110 A12226.

Tylka, J., C. M. S. Cohen, W. F. Dietrich, M. A. Lee, C. G. Maclennan, R. A. Mewaldt, C. K. Ng,, and D. V. Reames, Shock geometry, seed populations, and origin of variable elemental composition at high energies in large gradual particle events, The Astrophysical Journal, 625:474–495, 2005

Wang, Y., and I. Witten (1997). Inducing model trees for continuous classes. In Proceedings of Poster Papers, 9th European Conference on Machine Learning, 128-137, Prague.

Wei, W. (2005), Time Series Analysis, Univariate and Multivariate Methods, Pearson, Addison and Wesley.

Wichard, J. D. and M. Ogorzałek. (2007). Time series prediction with ensemble models applied to the CATS benchmark. Neurocomputing. 70, 13-15 (Aug. 2007), 2371-2378. DOI= http://dx.doi.org/10.1016/j.neucom.2005.12.136.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biometrics, 1, 80-8.

Xie, H., L. Ofman, and G. Lawrence (2004), Cone Model for Halo CMEs: Application to Space Weather Forecasting, *Journal of Geophysical Research*, 109, Issue A3.