Space Weather of the Heliosphere: Processes and Forecasts
Proceedings IAU Symposium No. 335, 2017
C. Foullon & O.E. Malandraki, eds.

doi:10.1017/S1743921317011012

Application of Test Particle Simulations to Solar Energetic Particle Forecasting

S. Dalla[1], **B. Swalwell**[1], **M. Battarbee**[1], **M. S. Marsh**[2], **T. Laitinen**[1] **and S. J. Proctor**[1]

[1]University of Central Lancashire, Preston, PR1 2HE, UK
[2]Met Office, Exeter, UK

Abstract. Modelling of Solar Energetic Particles (SEPs) is usually carried out by means of the 1D focused transport equation and the same approach is adopted within several SEP Space Weather forecasting frameworks. We present an alternative approach, based on test particle simulations, which naturally describes 3D particle propagation. The SPARX forecasting system is an example of how test particle simulations can be used in real time in a Space Weather context. SPARX is currently operational within the COMESEP Alert System. The performance of the system, which is triggered by detection of a solar flare of class >M1.0 is evaluated by comparing forecasts for flare events between 1997 and 2017 with actual SEP data from the GOES spacecraft.

Keywords. Sun: particle emission

1. The SPARX forecasting system

A number of physics-based forecasting models are currently being developed to predict the particle intensities and radiation risk associated with SEPs. The traditional approach is spatially 1D, meaning that particles are assumed to remain tied to the magnetic field line on which they were originally injected, with no propagation across the field (e.g. Aran *et al.* 2006; Luhmann *et al.* 2010). However, high energy SEPs have a strong influence on the radiation dose and these particles are affected by drifts due to the gradient and curvature of the Parker spiral, requiring a 3D description (Marsh *et al.* 2013; Dalla *et al.* 2013).

The SPARX forecasting system is based upon a 3D test particle model that accounts for drift and deceleration effects (Marsh *et al.* 2015). It is currently operational within the COMESEP alert system (www.comesep.eu, Crosby *et al.* 2012). In real time operation, SPARX is triggered by the detection of a solar flare of magnitude >M1.0. The system makes use of a database of runs of the test particle model to combine the contribution of many injection tiles near the Sun, describing an extended CME-driven shock structure. A profile of SEP intensities versus time at Earth for protons >10 MeV and >60 MeV is produced. Peak intensities are normalised by means of an empirical relationship between flare peak flux and particle peak intensities (Dierckxsens *et al.* 2015). Further details about SPARX may be found in (Marsh *et al.* 2015).

2. Evaluating model performance

To assess the performance of the SPARX model in forecasting SEP events and their parameters, we proceeded as follows. We started from a list of X-class solar flares that took place in the time range between 1 September 1997 and 30 April 2017, covering solar cycles 23 and 24 up to the present time. The initial list consisted of 169 flares.

hits = 20	false alarms = 27	47
misses = 20	correct negatives = 58	78
40	85	125

Table 1. Contingency table for SPARX forecasts for F_N threshold.

Bias	1.18
POD	0.5
FAR	0.57
POFD	0.32
CSI	0.30

Table 2. SPARX scores for F_N threshold.

For each flare, SPARX was run as it would have been in a forecast mode, i.e. with a fixed set of model input parameters which remained the same for all events. The output was then examined and a forecast constructed as follows: if the peak flux for protons >10 MeV over the entire duration of the event exceeded a specified threshold F, then a positive forecast of SEP event was made, otherwise a no event situation was predicted.

In the analysis below, we considered two thresholds: the first one is the standard NOAA threshold, F_N=10 pfu (where pfu is particles cm^{-2} s^{-1} sr^{-1}) and the second is a lower threshold F_1=1 pfu.

We then examined data from the >10 MeV proton channel of the GOES Energetic Particle Sensor (EPS) (Onsager *et al.* 1996) to verify whether an actual event took place following each of the flares in our list. To define the occurrence of an SEP event, we used in the first instance the threshold F_N and subsequently F_1. The methodology for associating an SEP event with a given flare is described in (Swalwell *et al.* 2017). At the start time of 36 flares in our list, the EPS detector already had an enhanced flux from a previous event, making it impossible to verify whether or not an SEP event occured: these flares were removed from the list. In addition, for 8 flares no EPS data was available. Therefore the total number of flares used in the analysis below is 125.

2.1. *NOAA threshold*

The contingency table for the NOAA threshold, F_N, is shown in Table 1. A correct negative is any X class flare which was not forecast to produce SEPs and which did not do so. From a contingency table, it is possible to derive a number of indicators that assess the performance of the model. They are defined as follows.

The bias score gives an indication of whether there is a tendency to under-forecast (bias < 1) or over-forecast (bias > 1) events. The perfect score is 1. The bias is given by:

$$\text{BIAS} = \frac{\text{hits} + \text{false alarms}}{\text{hits} + \text{misses}} \tag{2.1}$$

The probability of detection (POD) score gives the fraction of observed "yes" events which were correctly forecast. The perfect score is 1. The POD is given by:

$$\text{POD} = \frac{\text{hits}}{\text{hits} + \text{misses}} \tag{2.2}$$

The false alarm ratio (FAR) gives the fraction of predicted "yes" events which did not occur. The perfect score is 0. The FAR is given by:

https://doi.org/10.1017/S1743921317011012 Published online by Cambridge University Press

hits = 40	false alarms = 31	71
misses = 12	correct negatives = 42	54
52	73	125

Table 3. Contingency table for SPARX forecasts for F_1 threshold.

Bias	1.37
POD	0.77
FAR	0.44
POFD	0.43
CSI	0.48

Table 4. SPARX scores for F_1 threshold.

$$\text{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}} \tag{2.3}$$

The probability of false detection (POFD), also called false alarm rate, is the fraction of "no" events incorrectly forecast as "yes". The perfect score is 0. It is given by:

$$\text{POFD} = \frac{\text{false alarms}}{\text{correct negatives} + \text{false alarms}} \tag{2.4}$$

Finally the critical success index (CSI) shows how well the forecast "yes" events correspond to the observed "yes" events. The perfect score is 1. The CSI is given by:

$$\text{CSI} = \frac{\text{hits}}{\text{hits} + \text{misses} + \text{false alarms}} \tag{2.5}$$

For the SPARX model with F_N threshold, the values of the scores defined in eqs.(2.1)–(2.5) are given in Table 2.

2.2. *1 pfu threshold*

The contingency table is shown in Table 3 for the SPARX forecasts for the F_1 threshold. The various scores as defined in eqs.(2.1)–(2.5) are shown in Table 4 for the F_1 threshold.

3. Comparison between forecast and actual SEP profile characteristics

For those events which were correctly forecast by SPARX to occur, we can compare the properties of the forecast flux profiles with those of the observed GOES profiles. Figure 1 shows the peak fluxes forecast by the SPARX model versus actual peak fluxes from GOES EPS data. The colour scale of the symbols reflects the longitude of the flare. Figure 2 presents a comparison of times to maximum intensity, defined as the time between the start of the flare and the time of peak flux. SPARX times to maximum versus actual GOES EPS ones are plotted. As in Figure 1 the colour scale gives the flare longitude.

4. Discussion

SPARX performs much better if a lower threshold than that used by NOAA is used to define an SEP event. For both thresholds, the BIAS score is above 1, meaning that

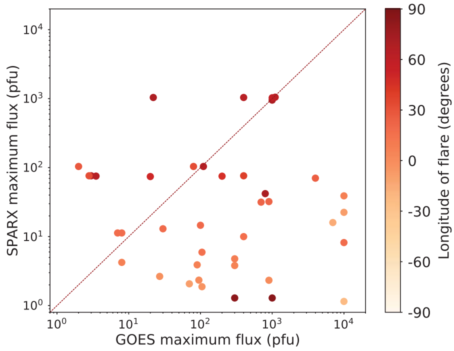

Figure 1. Peak fluxes forecast by the SPARX model versus actual GOES peak fluxes, for >10 MeV protons. The gray line is the 1:1 line.

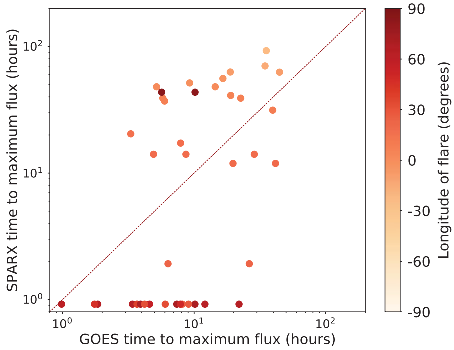

Figure 2. Time to peak intensity forecast by the SPARX model versus actual GOES time to peak intensity, for >10 MeV protons.

SPARX over-forecasts SEP events, more so for the lower threshold. All of the remaining scores are better for the lower threshold: the POD score for the lower threshold is much higher than for the lower (0.77 versus 0.5); the FAR is lower for the lower threshold (0.44 versus 0.57); and the CSI is higher for the lower threshold (0.48 versus 0.30).

Regarding the characteristics of the profiles of SEP fluxes, SPARX appears to significantly underestimate the time to peak intensity for well connected Western longitude flares, while for Eastern longitude flares reasonable agreement between forecast and actual times to peak are seen. Peak fluxes for Eastern longitude flares tend to be underestimated.

References

Aran, A., Sanahuja, B. & Lario, D. 2006, *Adv. Space Res.*, 37, 1240

Crosby, N. B., *et al.* 2012, in *AIP Conference Series*, vol. 1500, ed. Q. Hu, *et al.* p. 159

Dalla, S., Marsh, M. S., Kelly, J. & Laitinen, T. 2013, *J. Geophys. Res.*, 118, 5979

Dierckxsens, M., *et al.* 2015, *Solar Phys.*, 290, 841

Luhmann, J. G., *et al.* 2010, *Adv. Space Res.*, 46, 1

Marsh, M. S., Dalla, S., Kelly, J. & Laitinen, T. 2013, *Ap. J.*, 774, 4

Marsh, M. S., Dalla, S., Dierckxsens, M., Laitinen, T. & Crosby, N. B. 2015, *Space Wea.*, 13, 386

Onsager, T. G., *et al.* 1996, in *GOES-8 and Beyond, Proceedings of SPIE*, ed. E. R. Washwell, vol. 2812, p. 281

Swalwell, B., Dalla, S. & Walsh, R. W. 2017, *Solar Phys.*, 292, 173

https://doi.org/10.1017/S1743921317011012 Published online by Cambridge University Press