



A statistical study of the performance of the Hakamada-Akasofu-Fry version 2 numerical model in predicting solar shock arrival times at Earth during different phases of solar cycle 23

S. M. P. McKenna-Lawlor¹, C. D. Fry², M. Dryer^{2,3}, D. Heynderickx⁴, K. Kecskemety⁵, K. Kudela⁶, and J. Balaz^{1,6}

¹Space Technology Ireland, National University of Ireland, Maynooth, Co. Kildare, Ireland

²Exploration Physics International, Inc., Huntsville, Alabama, 35806, USA

³NOAA Space Environment Center, Boulder, Colorado, 80305, USA

⁴D-H Consultancy, Leuven, Belgium

⁵KFKI Research Institute for Particle and Nuclear Physics, Budapest, Hungary

⁶Institute of Experimental Physics, Kosice, Slovakia

Correspondence to: S. M. P. McKenna-Lawlor (stil@nuim.ie)

Received: 13 December 2010 – Revised: 20 February 2012 – Accepted: 21 February 2012 – Published: 27 February 2012

Abstract. The performance of the Hakamada Akasofu-Fry, version 2 (HAFv.2) numerical model, which provides predictions of solar shock arrival times at Earth, was subjected to a statistical study to investigate those solar/interplanetary circumstances under which the model performed well/poorly during key phases (rise/maximum/decay) of solar cycle 23. In addition to analyzing elements of the overall data set (584 selected events) associated with particular cycle phases, subsets were formed such that those events making up a particular sub-set showed common characteristics. The statistical significance of the results obtained using the various sets/subsets was generally very low and these results were not significant as compared with the hit by chance rate (50%). This implies a low level of confidence in the predictions of the model with no compelling result encouraging its use. However, the data suggested that the success rates of HAFv.2 were higher when the background solar wind speed at the time of shock initiation was relatively fast. Thus, in scenarios where the background solar wind speed is elevated and the calculated success rate significantly exceeds the rate by chance, the forecasts could provide potential value to the customer. With the composite statistics available for solar cycle 23, the calculated success rate at high solar wind speed, although clearly above 50%, was indicative rather than conclusive. The RMS error estimated for shock arrival times for every cycle phase and for the composite sample was in each case significantly better than would be expected for a

random data set. Also, the parameter “Probability of Detection, yes” (POD_y) which presents the Proportion of Yes observations that were correctly forecast (i.e. the ratio between the shocks correctly predicted and all the shocks observed), yielded values for the rise/maximum/decay phases of the cycle and using the composite sample of 0.85, 0.64, 0.79 and 0.77, respectively. The statistical results obtained through detailed analysis of the available data provided insights into how changing circumstances on the Sun and in interplanetary space can affect the performance of the model. Since shock arrival predictions are widely utilized in making commercially significant decisions re. protecting space assets, the present detailed archival studies can be useful in future operational decision making during solar cycle 24. It would be of added value in this context to use Briggs-Rupert methodology to estimate the cost to an operator of acting on an incorrect forecast.

Keywords. Interplanetary physics (Interplanetary shocks; Solar wind plasma)

1 Introduction

As ground based and space borne technological systems advance continuously in complexity, they correspondingly become more vulnerable to the particle radiation hazards posed by solar variability. The need to successfully predict the

arrival at Earth of solar related disturbances is thus becoming of ever increasing importance with time in both technological and scientific arenas.

Since environmental conditions related to shock transit can vary over the course of a solar cycle (e.g. due to variations in the complexity of the interplanetary background; changing helio-latitudes of flare initiation etc.) it is important that statistical studies extend over at least a full solar cycle in order to determine if such variations can affect the forecast outcome.

1.1 Predictive models

Several numerical models which use solar observations as inputs to forecast the arrival at Earth of flare related shocks have been already described in the literature and their relative performances discussed (e.g. McKenna-Lawlor et al., 2006, and references therein). These models include the “Shock Time of Arrival”/STOA Model (Dryer and Smart, 1984; Smart and Shea, 1984, 1985); the “Interplanetary Shock Propagation”/ISPM Model (Smith and Dryer, 1990); and the “Hakamada-Akasofu-Fry version 2”/modified HAFv.2 model (Dryer et al., 2001, 2004; Fry et al., 2001, 2003, 2004, 2005, 2007; McKenna-Lawlor et al., 2002, 2006; Smith et al., 2003, 2004, 2005; Sun et al., 2002a, b, 2003). Overviews of the characteristics of these three models are provided in Sect. 2.

More recently Tóth et al. (2005) developed a 3-D, numerical, full MHD model with the capability to simulate observed CMEs as they are driven outwards magnetically from active regions in the low corona. The methodology utilized involves a framework (termed the Space Weather Modeling Framework) which was developed by members of the Center for Space Environment Modeling (CSEM) at the University of Michigan. This framework acts to couple together several codes that individually model segments of the complex physical domain that extends from the corona to the Earth’s upper atmosphere and beyond.

The CSEM methodology couples: numerical models of: the solar corona; an eruptive event generator; the inner heliosphere; solar energetic particles; the global magnetosphere; the inner magnetosphere; the radiation belts; ionospheric electrodynamics and the upper atmosphere to provide a composite high-performance model which was initially employed to model the major coronal mass ejection (CME) of 28 October 2003 that formed part of the well known sequence of “Halloween events” generated in solar Active Region 10486 (Tóth et al., 2007; Manchester et al., 2008). The simulation created synthetic coronagraph images of the CME that were quantitatively compared with images recorded at L1 by the LASCO coronagraph aboard the SOHO spacecraft. These comparisons provided insights into the process of CME evolution that stimulated the development of improved global models of CME initiation. A method by Owens (2008) that combines remote and in situ

measurements of coronal mass ejections to model their large scale structure was initially used to investigate the geometry of the December 1996 Ulysses-SOHO quadrature event and further modeling is ongoing.

A coupling framework under development at the Center for Integrated Space Weather Modeling at Boston University, combines empirical, semi-empirical, and inverse models to construct a composite forecast model of the Sun-Earth system (Baker et al., 2004, 2007). See also methodologies that coupled (i) the coronal part of the Wang-Sheely-Arge model (Arge et al., 2004) with the ENLIL solar wind model (Odstrcil, 2003) and (ii) the coronal part of the MHD Around-A-Sphere model (Riley et al., 2001) with the ENLIL solar wind model to simulate the solar wind near the ecliptic plane at 1 AU (Lee et al., 2009). Preliminary testing of these ensembles involved comparing the predicted results with in situ measurements made at L1 during the declining phase of solar cycle 23.

Overall, steady progress is presently being made in modeling the Sun-Earth end-to-end system. In this connection the HAFv.2 model (Sect. 2) is currently being adapted to drive a 3-D MHD model by Detman et al. (2006) such that HAFv.2 is utilized to characterize the inner solar corona (from 2.5 R_s to 21.5 R_s) where the supersonic, super-Alfvénic, 3-D MHD model does not provide a solution.

The validation of various combinations of coupled codes is, at this time, a work in progress within the scientific community.

1.2 The present study

Shocks in the solar wind plasma and deformation of the ambient magnetic field associated with coronal mass ejections, stream-stream interactions and co-rotating interaction regions herald the occurrence of geomagnetic storms (Luhmann, 1997). Thus, the capability to forecast such events is critical to the establishment of successful space weather forecast capability.

A statistical study of circumstances under which the HAFv.2 model performed well/poorly in predicting shock arrivals at L1 during the rise and maximum phases of solar cycle 23 has already been carried out (McKenna-Lawlor et al., 2006) through subjecting to statistical analysis various event subsets within which the shocks recorded displayed common characteristics. In the present paper the data sample used is extended from the previous total of 339 events (February 1997–mid August 2002) to include a further 245 events recorded during the declining phase of the cycle (mid August 2002–December 2008). The performance of HAFv.2 in predicting shock arrival times at Earth in various subsets during this declining phase is investigated statistically and compared with corresponding data obtained during the previous two phases. Also, statistical results obtained through analyzing the total sample (584 events) are presented.

The data used when forming the various subsets were initially reported in three studies in which the individual shock arrival predictions of HAFv.2 were compared with in situ measurements made aboard spacecraft located at the L1 Lagrangian Point, which is located ~ 0.01 AU to sunward of Earth (i.e. at a distance less than the resolution provided by the HAFv.2 model). The periods covered by these studies were:

Study 1: February 1997–October 2000 during the cycle rise time (173 events) (Fry et al., 2003)

Study 2: November 2000–mid August 2002 during cycle maximum (166 events) (McKenna-Lawlor et al., 2006)

Study 3: Mid August 2002 through December 2008 during cycle decay (245 events) (Smith et al., 2009)

The interested reader will find in these papers details of the individual shock events recorded during each cycle phase.

The methodology for measuring shock arrival times in data recorded by the “Solar Wind Experiment Proton Alpha Monitor” (SWEPAM) and the “Magnetometer Experiment” (MAG) aboard the “Advanced Composition Explorer” (ACE) spacecraft at L1 is already described in McKenna-Lawlor et al. (2006). It is noted that the shocks thereby identified in the ACE data were individually checked against complementary shock events measured at L1 by the Solar Heliospheric Observatory/SOHO spacecraft experimenters (available at <http://umtof.umd.edu/pm/>).

The shock events sourced from the three studies mentioned include only those for which coincident (within $1/2$ h) radio and X-ray data were available in near real time among global observations reported by individual experimenters to the forecasters at the “Space Environment Centre” at Boulder, Colorado. Events involving estimated values were discarded from these data. The procedure used in selecting each of the individual shock events analyzed is not described here since this matter is already covered in detail by the authors of the three studies mentioned above. It is, however, worth pointing out that, since the same forecasters were involved in event selection during all three phases of solar cycle 23, the overall data set is a particularly uniform one.

1.3 Present study outline

Section 2 provides characteristics of the STOA, ISPM and HAFv.2 models. Also, a brief description is provided of how HAFv.2 is used to predict the arrival at Earth of interplanetary shocks in a “near real time” operational environment. Section 3 presents terminology originally developed by the metrological community to describe the success, or otherwise, of modeled weather predictions. These same terms are utilized in the present study to classify the performance of HAFv.2 shock predictions. Various parameters and skill scores used to evaluate the predictive results of the HAFv.2 model are introduced and an account presented of how the

results obtained through statistical analysis can be evaluated in terms of a “Figure of Merit” (sr) and their statistical significance expressed using a χ^2 test with significance levels for three degrees of freedom. Next, the criteria applied when forming subsets from the composite data set, such that the events making up particular sub-sets show common characteristics, are indicated and a list of the subsets selected for investigation provided.

In Sect. 4 the results obtained through subjecting the various sets and subsets defined in Sect. 3 to statistical analysis are presented and considered. In Sect. 5 aspects of the statistical results obtained are discussed. Also it is recommended that users of predictive results employ Briggs–Rupert (BR) methodology to calculate for individual predictions error estimates that indicate the cost to that user of acting on an incorrect forecast. The possible influence of changing solar/interplanetary conditions on the model predictions at different cycle phases is also considered. Conclusions are presented in Sect. 6.

2 Outline of the STOA, ISPM and HAFv.2 models

The STOA and ISPM models each require as input data the initial coronal shock velocity, as well as the piston driving time and location on the Sun of associated flares. STOA utilizes the solar wind velocity measured in situ at 1 AU (available in real time from the National Oceanic and Atmospheric Administration at Boulder) to estimate the “Parker-type”, spherically symmetric and polytropic solar wind velocity profile out to 1 AU, assuming that a uniform solar wind in heliolongitude is present upstream of the shock. If such measured data are not available a default value of 400 km s^{-1} is adopted. ISPM, on the other hand, is based on a single background solar wind model with a representative solar wind speed of 360 km s^{-1} at 1 AU. Neither STOA nor ISPM takes into account possible stream-stream interactions in Sun–Earth space. Thus, if a prior shock event has occurred within 24 h the predictions of the models are rendered uncertain, due to the possibility that temporally close solar wind events may interact.

Shock Arrival Times (SATs) are estimated by STOA through computing predicted solar wind speed, density and dynamic pressure for several future days. These temporal profiles of simulated shock speed, relative to the ambient, Parker-type, solar wind speed, are then scanned automatically to calculate the magneto-acoustic Mach Number (Ma) at the Earth’s position, using a representative value of the interplanetary magnetic field (IMF). Values of $\text{Ma} < 1$ indicate that the shock has decayed to an MHD wave.

SATs are determined by ISPM through again making advance computations of the solar wind speed, density and dynamic pressure. The simulated dynamic pressure is then scanned automatically to calculate a shock strength index/SSI such that $\text{SSI} = \log(\Delta P/P_{\min})$, where P is the

dynamic pressure; ΔP is the difference in pressure during consecutive time steps and P_{\min} is the minimum pressure for these time steps. $SSI=0$ represents a threshold value equivalent to the limit $Ma=1$ below which shocks decay to an MHD wave (Dryer, 1974).

HAFv.2 uses the same observational inputs as STOA and ISPM but differs from these models in the way the background solar wind is treated. HAFv.2 utilizes a model of the inhomogeneous, ambient, solar wind that affects the propagation of disturbances from the Sun to Earth. Realistic inner boundary conditions are used to determine the background solar wind flow, as well as the relevant IMF morphology (see below). The model is particularly useful in that it provides, based on its input solar parameters, information concerning the non-uniform conditions prevailing in the heliosphere through which particular solar shocks propagate (McKenna-Lawlor et al., 2002, 2006). In addition, it monitors how the point which links the interplanetary magnetic field/IMF lines with the shock (which is called the COB/Connection with OBServer point following Lario et al., 1998), changes its location progressively, thereby influencing the local rise times of individual temporal flux enhancements.

HAFv.2 presents an advantage over STOA and ISPM in that, through estimating solar wind propagation in a realistic solar wind matrix, it can indicate if stream-stream interactions were present upstream of Earth (Fry et al., 2001). Also, it provides a means of distinguishing solar event driven shocks from those generated by the passage of co-rotating interaction regions (McKenna-Lawlor et al., 2006). These advantages have prompted its present use for predictive purposes by the United States Airforce although its predictive capability is not significantly better than that of STOA and ISPM.

2.1 Application of HAFv.2 in an operational environment

HAFv2 comprises a 3-D code which models the non-uniform flow of the solar wind, the IMF and the plasma speed and plasma density as fluid parcels that are projected outward from a Sun-centered, spherical, rotating surface located at 2.5 Rs. The model is kinematic in that it kinetically projects the flow of the solar wind from inhomogeneous sources near the Sun outward into interplanetary space. It is “modified” in that it adjusts the flow for stream-stream interactions as faster plasma catches up with slower plasma. Potential Field Source Surface maps available from the *Space Environment Centre* provide solar wind speed and radial magnetic field at the HAFv.2 inner boundary which is set at 2.5 Rs (Wang and Sheeley, 1990; Arge and Pizzo, 2000). These data are continuously updated (website: <http://www.sec.noaa.gov>) to reflect ongoing changes in solar circumstances, so that the background solar wind can be estimated at specific user locations having regard to ongoing changes at the inner boundary of the model.

Transient events are modeled by superposing pulses on the prevailing (background) solar wind. The input parameters associatively used are:

- Optical/X-ray event start time (within up to 0.5 h of the accompanying shock start)
- Disk location of the parent solar event
- Event duration (proxy piston driving time of shock: determined from the GOES soft X-ray profile of the flare)
- Shock start (determined from metric Type II radio burst data)
- Initial speed (V_s) of the shock near the Sun (estimated from reported metric Type II speed, or plane of the sky CME speed)

Information concerning the limitations inherent in the use of these various kinds of data is contained in McKenna-Lawlor et al. (2006).

Predicted shock arrival times are extracted from automatic scans of the temporal profiles of the simulated terrestrial dynamic pressure using a Shock Search Index (SSI) such that: $SSI = \log(\Delta P/P_{\min})$, where P is either the dynamic pressure or the momentum flux; ΔP is the difference in P during consecutive 1 h time steps, and P_{\min} is the minimum P value for these time steps. $SSI=0$ represents a threshold value equivalent to the limit $Ma=1$ below which shocks decay to an MHD wave. The threshold adopted can be pre-set to provide an optimum value for a particular phase of the solar cycle.

HAFv.2 provides the solar wind speed, density, dynamic pressure and IMF vector at selected locations in the heliosphere (in the present instance at Earth). For further details see Hakamada and Akasofu (1982), Sun et al. (1985) and Fry et al. (2001). It is noted that the model has also shown success in predicting shock arrivals at Mars and Venus (McKenna-Lawlor et al., 2005, 2008) and this predictive capability is presently being tested using improved statistics.

3 Classification of the success of predicted shock arrival times

The success of predicting shock arrival times at L1 can be expressed through adopting standard meteorological terminology (Schaefer, 1990). In the present text the following fundamental definitions described by Smith et al. (2009) are adopted in classifying the Hits (h), False Alarms (fa), Misses (ms, mf) and Correct Nulls (cn, cn/ int) in the present sample of 584 events.

Table 1. Definitions of various meteorological standard forecast skill scores used to evaluate the performance of numerical predictive models. These constitute various combinations of the contingency values and the specific parameter of interest to a user depends on the needs of that particular user.

Statistic			Definition
Probability of detection, Yes	PODy	$a/(a+c)$	Proportion of Yes observations that were correctly forecast
Probability of detection, No	PODn	$d/(b+d)$	Proportion of No observations that were correctly forecast
False alarm ratio	FAR	$b/(a+b)$	Proportion of Yes forecasts that were incorrect
Bias	BIAS	$(a+b)/(a+c)$	Ratio of number of Yes forecasts to Yes observations
Critical success index	CSI	$a/(a+b+c)$	Proportion of hits that were either forecast or observed
True skill statistic	TSS	$(\text{PODy} + \text{PODn} - 1)$	Measures ability to discriminate between Yes and No
Heidke skill score	HSS	$(a+d-C1)/(N-C1)$	% correct, corrected by those expected correct by chance
Gilbert skill score	GSS	$(a-C2)/(a+b+c-C2)$	CSI, corrected by number of hits expected by chance (C2)

where $a = h$; $b = \text{fa}$; $c = \text{m}$; $d = \text{cn}$; $N = (a+b+c+d)$ is the total number of events; $C1 = C2 + (b+d)(c+d)/N$ is the number of events (h+cn) expected to be correct by chance. $C2 = (a+c)(a+b)/N$ is the number of hits expected to occur by chance.

Hit (h):	Shock observed within ± 24 h of the predicted shock.
False Alarm (fa):	Shock predicted but not observed within 1–5 days of the solar event.
Miss (ms):	A shock not predicted but observed within 1–5 days of the solar event (i.e. it was preceded by a solar event but not predicted properly).
Miss/fa (mf):	A shock both predicted and observed but not within 24 h of each other. These can be indicated by both “miss” and “false alarm” but are added to the miss category when computing statistics.
Correct Null (cn):	No shock predicted and none detected within 1–5 days of the solar event.
cn/int:	A correct null due to interaction with another shock. No separate shock is predicted. These events are differentiated from the “cn” category in order to show events that are part of enhanced solar activity but they are added to the “cn” list when computing statistics.

A 24-h window is adopted as a general working parameter. This and other possible hit windows were discussed earlier in a HAFv.2 context by Mozer and Briggs (2003). The effect of increasing/decreasing the size of the hit window was also considered by McKenna-Lawlor et al. (2006) and it was

demonstrated by the latter authors that the success rate of HAFv.2 was substantially lower when the hit window was reduced to 12 h. In addition it was demonstrated that the hit window should not be extended above 36 h since, at very wide time intervals, some of the events formerly classed as misses become hits.

The statistical parameters used to evaluate the performance of a predictive model may be derived through forming various combinations of the basic hit/false alarm/miss/correct null categories. For convenience these are individually assigned contingency values a , b , c , d [a = number of hits (forecast yes, observed yes); b = false alarms (forecast yes, observed no); c = misses and d = correct nulls (good predictions, forecast no, observed no)] for events in a particular data set N (where $N = a + b + c + d$).

Parameters used to evaluate the predictive results of the HAFv.2 model in association with different solar cycle phases are defined in Table 1. These standard skill scores are composed of various combinations of the contingency values and the specific parameter of interest to a user depends on the individual needs of that user.

For instance, the Probability of Detection, yes (PODy) parameter gives the ratio between the shocks correctly predicted and all the shocks observed. Thus a value of 1 indicates that all the observed shocks were correctly forecast.

If a user is in particular interested in avoiding false alarms, the “Probability of Detection, no” (PODn) parameter (Table 1) which compares the numbers of correct nulls and false alarms is of special importance. A value of $\text{PODn} = 1$ means that the predictions were completely successful (no false alarms and at least one correct null). A value of $\text{PODn} = 0$ indicates that there were no correct nulls and at least one false alarm. For the composite sample PODn was 0.41.

The “False Alarm Ratio”, FAR, compares false alarms with hits. Here the prediction would be fully successful (no false alarms and at least one hit) if $FAR=0$, whereas FAR approaching 1 would indicate that there were fewer hits than false alarms.

The standard skill scores are useful as broad indicators of the performance of a forecast system. However, it is pointed out that among the various standard parameters listed in Table 1 only PODn, the “True Skill Statistic” TSS and the “Heidke Skill Score” HSS involve correct nulls. An additional measure is thus introduced to describe the success rates of the predictions relative to the measurements. This percentage success rate in forecasting (sr), which is sometimes referred to as the “Figure of Merit”, is defined by:

$$sr = [(h + cn) \times 100 / N], \quad (1)$$

A χ^2 test may be applied to the result such that:

$$\chi^2 = \Sigma(O - E)^2 / E, \quad (2)$$

where O represents the numbers contained in the event tables, and E is estimated using the formulae:

$$\text{for } h: (a + c)(a + b) / N \quad (3)$$

$$\text{for } fa: (a + b)(b + d) / N \quad (4)$$

$$\text{for } m: (a + c)(c + d) / N \quad (5)$$

$$\text{for } cn: (b + d)(c + d) / N \quad (6)$$

It is noted that $(O - E)$ is the same for all four quantities and is given by $(ad - bc) / N$. P values are estimated in the present case for three degrees of freedom. A value of $p_3 < 0.05$ indicates a high level of significance while $0.05 < p_3 < 0.2$ indicates a lesser, but still acceptable, level of significance.

3.1 Formation of data subsets

Since the overall data set includes shocks that were generated during different phases of solar cycle 23 and thus subsequently travelled to Earth under different interplanetary conditions, sub-sets were constructed such that the events making up a particular sub-set showed common characteristics.

The subsets selected for investigation comprise shocks events associated with

- Minor flares (subflares)
- Flares with short ($\tau \leq 20$ m) proxy piston driving times
- Events with solar wind background $>400 \text{ km s}^{-1}$ and $\leq 400 \text{ km s}^{-1}$
- Flares originating at heliolongitudes $>20^\circ$ and $\leq 20^\circ$
- Shocks with speeds $<1200 \text{ km s}^{-1}$ at heliolongitudes $>20^\circ$
- Shocks with speeds $<1200 \text{ km s}^{-1}$ at heliolongitudes $\geq 80^\circ$

Table 2. Overview of the values of a range of statistical parameters derived using the HAFv.2 model for the key phases of solar cycle 23 and for the total cycle.

	All	Rise	Max	Decl
Probability of detection, yes (PODy)	0.77	0.85	0.64	0.79
Probability of detection, no (PODn)	0.41	0.32	0.45	0.44
False alarm ratio (FAR)	0.54	0.55	0.64	0.48
Bias (BIAS)	1.69	1.90	1.75	1.53
Critical success index (CSI)	0.40	0.42	0.30	0.45
True skill statistic (TSS)	0.18	0.18	0.09	0.23
Heidke skill score (HSS)	0.16	0.16	0.07	0.22
Gilbert skill score (GSS)	0.09	0.09	0.04	0.12
Number in each set (N)	584	173	166	245
Root mean square of DT, hours	11.4	11.6	11.4	10.6
Average of DT , hours	9.3	9.6	9.2	9.2
% correct = $[(h + cn) \cdot 100] / N$	55 %	53 %	51 %	59 %
% of events followed by observed shock = $100 \cdot (h + m) / N$	39 %	39 %	33 %	43 %
% of observed shocks well-predicted = $100 \cdot h / (h + m)$	77 %	85 %	64 %	79 %

4 Results of the statistical analysis of data recorded during the rise/max/decay phases of solar cycle 23 as well as using the composite sample

An overview of values obtained using the HAFv.2 model for a range of statistical parameters estimated over the key phases of solar cycle 23 is presented in Table 2. The right hand column gives the values for the declining phase of the cycle and these are compared in the table with corresponding values for the maximum and rise phases as well as for the composite sample.

These verification statistics are of potential value to users making commercial decisions based on new predictions obtained during solar cycle 24. It is shown in Table 2 that the parameter “Probability of Detection, yes” (PODy) which presents the Proportion of Yes observations that were correctly forecast (i.e. the ratio between the shocks correctly predicted and all the shocks observed, so that a value of 1 indicates that all the observed shocks were correctly predicted) yielded a value during the cycle rise phase of 0.85. During the maximum phase it was 0.64 and during the declining phase 0.79. For the composite sample the corresponding value was 0.77.

In Table 3, summary data concerning the performance of HAFv.2 with regard to 245 events recorded during the decay phase of solar cycle 23 are compared with complementary data reported by McKenna-Lawlor et al. (2006) for the rise phase (173 events) and maximum phase (166). Data for the composite sample (584 events) are also shown.

Table 3. Summary information concerning the performance of the HAFv.2 model with respect to data measured during the key (rise/max/decline) phases of solar cycle 23 and over the total cycle.

Status	Nr. of events	(h)	(fa)	(cn)	(m)	(h/m)	(sr)	χ^2	p_3
Cycle rise	173	58	71	34	10	5.8	53	6.80	0.009
Cycle max	166	35	61	50	20	1.7	51	1.14	0.286
Cycle decay	245	83	78	62	22	3.8	59	14.50	0.0001
Composite	584	176	210	146	52	3.4	55	20.55	<0.0001

It is seen that a relatively large number of events were recorded during the decay phase and that the hit rate was greatest at that time. The (sr) values for the rise/max/decay phases and for the composite sample were 53 %, 51 %, 59 % and 55 %, respectively. The results obtained for the rise and decay phases and for the composite sample are of high statistical significance. The result obtained during cycle maximum is marginally of statistical significance (see the discussion in Sect. 5.2).

In what follows, statistical results obtained through analyzing the subsets described in Sect. 3.2 are presented.

4.1 Statistics of minor flare events

Flare data used in the present study were taken from reports of X-ray image measurements made by the “Soft X-ray Imager” (SXI) aboard the GOES spacecraft. In this list, X-ray flare intensity is expressed in terms of the letters A, B, C, M and X, where $A = 1.0 \times 10^{-8}$ (W m^{-2}); $B = 1.0 \times 10^{-7}$ (W m^{-2}); $C = 1.0 \times 10^{-6}$ (W m^{-2}); $M = 1.0 \times 10^{-5}$ (W m^{-2}) and $X = 1.0 \times 10^{-4}$ (W m^{-2}). In addition, each letter has nine subdivisions (C1–C9).

Optical flares (recorded routinely by the United States Air Force SOON stations and reported to the Space Environment Services Center at Boulder) are classified in another way. Importance 0 (sub-flares are usually designated by S) and display an area ≤ 2 hemispheric square degrees while importance 1 flares have areas in the range: 2.1–5.1 square degrees (where one square degree = 1.214×10^4 $\text{km}^2 = 48.5$ millionths of the visible solar hemisphere). A brightness qualifier (F – faint, N – normal and B – brilliant) is generally added to the number that indicates area.

As noted by McKenna-Lawlor et al. (2006), optical events of class SF can be variously associated in the reported data with X-ray events of classes C, M and X. Geometrical effects at the limb and poor seeing contribute to the differences concerned. In view of this inherent variability, the data set is not tested here for the effect of optical sub-flares but only with respect to minor flares of X-ray classes C1–C9.

Table 4 provides an overview of the performance of the HAFv.2 model with respect to predicting SATs associated with minor X-ray flares during the key phases of solar cycle 23, and also over the complete cycle.

It is seen that very similar numbers of minor flares occurred during each cycle phase. Also, that each of these subsets was associated with a large number of false alarms. Less than half the number of hits were scored during cycle maximum than were scored during cycle rise and decay. Also, at the time of cycle maximum the number of misses was relatively high. This latter result may reflect the inhomogeneous character of the interplanetary medium at the time of cycle maximum.

The data indicate overall that events that are relatively minor at X-ray wavelengths cannot be presupposed not to produce a shock at L1 and thus their potential contribution cannot be neglected. This is in accord with an earlier statement by Fry et al. (2003) based on rise phase data that, while the HAFv.2 false alarm rate could be reduced by routinely predicting “no shock” for minor optical flare events (classes SF, SN, SB and 1N), a number of “hits” then became “misses”, thereby degrading the forecast skill scores already obtained. In these circumstances it was recommended to the forecasters not to use flare classification as a basis for yes/no shock arrival decisions.

The statistical level of the results obtained in this category was low overall. The figure of merit for the composite sample was $sr = 53$ % with $p_3 = 0.1392$ and it is noted that the large number of correct nulls identified contributed to producing this result.

4.2 The statistics of short ($\tau \leq 20$ m) proxy piston driving times

In the HAFv.2 model, the temporal profile of the shock speed at the Sun is considered to be governed by a time constant (τ) that appears in the exponential expression:

$$V(t) = V_s(t/\tau)\exp(1 - t/\tau) \quad (7)$$

where the τ value is a proxy diagnostic, determined from the integrated X-ray flux in the 1–8 Å channel of the GOES spacecraft (Akasofu, 2001). This parameter is estimated, on a logarithmic flux scale, to be the time duration measured at one-half the distance from the pre-event background level to the peak. The shock speed rises exponentially to an assumed maximum (V_s) and, thereafter, falls to a final decayed value of the background solar wind speed.

Table 4. Performance of the HAFv.2 model with respect to the prediction of SATs in association with a population of minor flares of X-ray classes C1-C9 during the key phases of solar cycle 23 and during the total cycle.

Status	Nr. of events	(h)	(fa)	(cn)	(m)	(h/m)	(sr)	χ^2	p_3
Cycle rise	61	15	29	15	2	7.5	49	3.05	0.3840
Cycle max	69	6	22	33	8	0.7	57	0.04	0.9979
Cycle decay	67	17	28	18	4	4.3	52	2.64	0.4505
Composite	197	38	79	66	14	2.7	53	5.49	0.1392

Table 5. Performance of the HAFv.2 model with respect to predicting SATs in association with a population of shocks with short ($\tau \leq 20$ min) proxy piston driving times during the key phases of solar cycle 23 and during the total cycle.

Status	Nr. of events	(h)	(fa)	(cn)	(m)	(h/m)	(sr)	χ^2	p_3
Cycle rise	41	9	20	11	1	9.0	49	2.38	0.4974
Cycle max	54	4	23	20	7	0.6	44	1.03	0.7940
Cycle decay	33	4	16	11	2	2.0	45	0.11	0.9906
Composite	128	17	59	42	10	1.7	46	0.18	0.9808

In the light of the relatively frequent occurrence of shock events featuring very short (proxy) driving times (τ), a subset was formed to include cases where $\tau \leq 20$ min. Table 5 shows the performance of the HAFv.2 model with respect to these events during the key phases of solar cycle 23 and during the overall cycle.

These data indicate that if such events were excluded there would be a substantial reduction in the successful prediction of hits by HAFv.2, particularly during the cycle rise phase, and it cannot be assumed if $\tau \leq 20$ min. for a particular event that no shock will follow. It is, however, noted that the number of false alarms substantially exceeded the hit rate in each cycle phase. Also, that there were substantially fewer events (33) with short driving times in the decay phase than during the rise (41) and maximum phases (54), with the largest number of events recorded during cycle maximum. This may reflect the nature of the flaring produced as the cycle evolved. Further, the fact that more than twice as many hits were registered during the rise time than during other cycle phases suggests that the changing state of complexity of the interplanetary medium as the cycle progressed may have played a role in the overall outcome.

The (sr) values obtained for the key phases of the cycle and for the composite sample were low (49 %, 44 %, 45 % and 46 %, respectively) and application of a χ^2 test to the predictions indicates that the results obtained in this category were not statistically significant. It is pointed out here that, in the modeling process, the power exerted is assumed in each case to be continuous throughout τ (i.e. a piston driven shock is assumed, before it decays, to be impelled away from the Sun throughout τ at its initially measured Type II speed). This is presently only an assumption and it may be that lack of justification of this hypothesis is responsible for the poor result obtained in this category.

4.3 Statistics of the background solar wind speed

To determine if the speed of the background solar wind affected model performance, shocks recorded during solar cycle 23 were divided between those for which the background solar wind speed was $>400 \text{ km s}^{-1}$ at the time of shock initiation and those for which it was $\leq 400 \text{ km s}^{-1}$ (Table 6, top/bottom).

Table 6 (top) shows that twice as many shocks (180) were recorded during solar cycle decay than were detected during the other two cycle phases (80, 82) when the solar wind was relatively fast. The performance of the model during the decay phase under elevated solar wind conditions was characterized by an (sr) rate of 62 % and showed a high level of statistical significance due to the accompanying high hit (h) and correct null (cn) rates, although there was an accompanying significant rise in the false alarm (fa) rate. The success rate (sr) was ≥ 55 % during each of the cycle phases. The result obtained using the composite sample showed, with high significance, that relatively fast solar wind speed was characterized by a 58 % success rate.

When, on the other hand, the solar wind speed was $\leq 400 \text{ km s}^{-1}$ (Table 6, bottom) the number of events recorded during solar cycle decay (63) was substantially lower than was the case during its complementary rise and maximum phases (93, 84). The success rate of the model in these circumstances was of the order of 50 % while the statistical significance of the results obtained was very low overall.

These data indicate that the success rates of HAFv.2 were higher when the background solar wind speed at the time of shock initiation was relatively fast.

Table 6. Performance of the HAFv.2 model with respect to the speed of the background solar wind during the key phases of solar cycle 23 and during the total cycle. The top four panels show the statistics when the speed was $>400 \text{ km s}^{-1}$ and the bottom four panels when it was $\leq 400 \text{ km s}^{-1}$ at the times of shock initiation.

Status	Nr. of events	(h)	(fa)	(cn)	(m)	(h/m)	(sr)	χ^2	p_3
Solar Wind Speed $>400 \text{ km s}^{-1}$									
Cycle rise	80	26	31	18	5	5.2	55	3.94	0.2680
Cycle max	82	20	30	25	7	2.9	55	2.90	0.4073
Cycle decay	180	68	56	43	13	5.2	62	15.59	0.0014
Composite	342	114	117	86	25	4.6	58	22.36	<0.0001
Solar Wind Speed $\leq 400 \text{ km s}^{-1}$									
Cycle rise	93	32	40	16	5	6.4	52	2.88	0.4105
Cycle max	84	15	31	25	13	1.2	48	0.02	0.9993
Cycle decay	63	14	21	19	9	1.6	52	0.41	0.9382
Composite	240	61	92	60	27	2.3	50	1.86	0.6020

Table 7. Performance of the HAFv.2 model with respect to shocks associated with flares located at helio-longitudes $>20^\circ$ (top four panels) and at heliolongitudes $\leq 20^\circ$ (bottom four panels) during each of the key phases of solar cycle 23 and during the total cycle.

Status	Nr. of events	(h)	(fa)	(cn)	(m)	(h/m)	(sr)	χ^2	p_3
Solar longitude $>20^\circ$									
Cycle rise	125	32	55	32	6	5.3	51	5.50	0.1386
Cycle max	117	24	36	44	13	1.8	58	3.40	0.3340
Cycle decay	182	55	56	51	20	2.8	58	8.17	0.0426
Composite	424	111	147	127	39	2.8	56	16.86	0.0008
Solar longitude $\leq 20^\circ$									
Cycle rise	48	26	16	2	4	6.5	58	0.05	0.9971
Cycle max	49	11	25	6	7	1.6	35	2.22	0.5280
Cycle decay	63	28	22	11	2	14.0	62	6.82	0.0779
Composite	160	65	63	19	13	5.0	53	1.06	0.7867

4.4 Statistics of shock arrivals from flares originating at heliolongitudes $>20^\circ$ and $\leq 20^\circ$

The performance of HAFv.2 with respect to shocks originating from flares at heliolongitudes $>20^\circ$ and $\leq 20^\circ$ was next investigated (Table 7).

These data show that a substantially higher number of measured events originated at high than at low helio-longitudes (424 vs. 160) during solar cycle 23. In each case a preponderance of events occurred during the decay phase. During this latter phase, although there was in each instance a large number of false alarms, the success rate was 58 % at high solar longitudes and 62 % at low solar longitudes, each with a high level of statistical significance.

4.5 Statistics of shocks with speeds $<1200 \text{ km s}^{-1}$ at helio-longitudes $>20^\circ$

Since the results provided in Table 7 were noted to pertain to a mixed population of fast and slow shocks, the effect of shock speed on the statistical outcome was next investigated. Table 8, presents statistical results concerning the

number of shocks with speeds $<1200 \text{ km s}^{-1}$ originating at helio-longitudes $>20^\circ$ during the key phases of the cycle and over the full cycle. Those high speed shocks recorded close to central meridian ($<20^\circ$) were found not to be suitable for statistical analysis due to their very low number and they are not considered further here.

Table 8 shows that the number of hits recorded for relatively low speed shocks during the rise and decay phases of the solar cycle were greater by a factor of two than was the case at solar maximum. The cn rate substantially exceeded the hit rate in all cycle phases. This category was associated with a high number of false alarms. Predictions associated with the composite sample (327 events) were characterized by an (sr) value of 56 % and the result was statistically significant, primarily due to the high cn rate.

4.6 Statistics of shocks with speeds $<1200 \text{ km s}^{-1}$ at heliolongitudes $\geq 80^\circ$

During the rise phase of the cycle shock speeds in the sample ranged between $500\text{--}1400 \text{ km s}^{-1}$ with 86 % displaying speeds $<1000 \text{ km s}^{-1}$. At solar maximum the speed range

Table 8. Performance of the HAFv.2 model with respect to shocks with speeds $<1200 \text{ km s}^{-1}$ at helio-longitudes $>20^\circ$ during each of the key phases of solar cycle 23 and during the total cycle.

Status	Nr. of events	(h)	(fa)	(cn)	(m)	(h/m)	(sr)	χ^2	p_3
Cycle rise	109	28	45	32	4	7.0	51	8.63	0.0346
Cycle max	95	13	31	42	9	1.4	58	1.87	0.5998
Cycle decay	123	26	38	41	18	1.4	54	1.37	0.7126
Composite	327	67	114	115	31	2.2	56	9.60	0.0223

Table 9. Performance of the HAFv.2 model with respect to shocks with speeds $<1200 \text{ km s}^{-1}$ originating at helio-longitudes $>80^\circ$ during each of the key phases of solar cycle 23 and during the total cycle.

Status	Nr. of events	(h)	(fa)	(cn)	(m)	(h/m)	(sr)	χ^2	p_3
Cycle rise	21	5	5	9	2	2.5	66	2.40	0.4936
Cycle max	21	1	7	10	3	0.3	52	0.35	0.9504
Cycle decay	47	13	16	15	3	4.3	60	3.93	0.2691
Composite	89	19	28	34	8	2.4	60	4.79	0.1878

spanned $398\text{--}2900 \text{ km s}^{-1}$ with $\sim 57\%$ displaying speeds $<1000 \text{ km s}^{-1}$. In the decay phase the speeds spanned $325\text{--}3000 \text{ km s}^{-1}$ with 55% showing speeds $<1000 \text{ km s}^{-1}$.

Since a high proportion of the shocks arriving at the Earth from limb events were relatively slow, particularly during the rise phase of the cycle, these events were selected for analysis (uncertainties in measured shock velocities due to horizontal variations in the coronal density gradient, combined with the non-radial shock propagation observed from Earth, can be expected to result in a better success rate for this population than for fast shocks).

A subset of 89 shock events with speeds $<1200 \text{ km s}^{-1}$ that originated in association with flares at helio-longitudes $\geq 80^\circ$ was identified. This sample was made up of 21 events during both the rise and maximum phase of the cycle while more than twice as many such events (47) occurred during the decay phase. The hits were associated with both east-limb and west-limb flares and their parent shocks covered a wide range of speeds, rather than being uniformly fast.

The performance of HAFv.2 in predicting shocks with speeds $<1200 \text{ km s}^{-1}$ originating at helio-longitudes $>80^\circ$ had an sr value of 60 % during the decay phase of the cycle but this result showed marginal statistical significance. The composite (89 events) sample was also characterized by an sr value of 60 % and it had a p_3 value of 0.1878.

5 Discussion

5.1 Overview of the statistical results

The information presented in Tables 2–9 express in different ways the success of HAFv.2 in estimating, under different solar/interplanetary circumstances, shock arrival times at

Earth. With reference to the statistics concerned, it is pointed out that, in an earlier paper (McKenna-Lawlor et al., 2006), the transit times predicted by “sample climatology” were estimated. In this regard, for each solar event a “sample climatology” prediction (TTc) was taken to be the average of all the preceding events for which there was an observed transit time (TTd). If the immediately preceding solar event did not have an associated shock at 1 AU, then TTc was assigned the value of the event that occurred prior to the immediately preceding event. “True climatology” is defined to comprise this same kind of information determined over a full cycle.

Since “true climatology” cannot predict complete nulls, it was considered that it would not be useful for the present paper to calculate the figures it provides and compare them with the results obtained by HAFv.2. Instead, the Root Mean Square error (RMS) for ΔT (i.e. the predicted minus the observed shock arrival time) for each phase of the solar cycle was estimated. For a random data set of shock arrival times evenly spread throughout the $\pm 24 \text{ h}$ window utilized a RMS error of 14.2 h is expected during the rise phase of the cycle (Fry et al., 2003). The RMS value yielded by HAFv.2 for this phase of the cycle was 11.6. In the maximum and declining phases of the cycle the corresponding RMS values were 11.4 and 10.6, respectively. For the composite sample the RMS value was 11.4. The values obtained using HAFv.2 were thus in every case, better than would be expected for a random set.

The statistical significance of the results obtained using the various sets/subsets was generally very low and these results were not significant as compared with the hit by chance rate (50 %), thus providing a low level of confidence in the predictions of the model with no compelling result encouraging its use. However, the data indicated that the success rates of

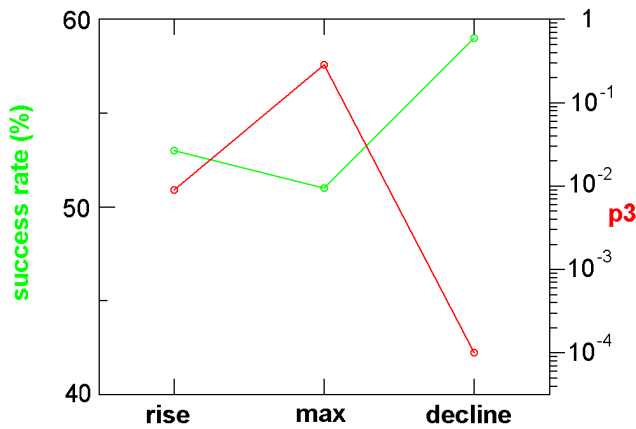


Fig. 1. The success rate *sr* (in green, left co-ordinate) and p_3 (in red, right co-ordinate) plotted for the events measured during each of the three phases of solar cycle 23 (see Table 3).

HAFv.2 were higher when the background solar wind speed at the time of shock initiation was relatively fast. Thus, in scenarios where the background solar wind speed is elevated and the calculated success rate significantly exceeds the hit rate by chance, the forecasts could provide potential value to the customer. With the composite statistics available for solar cycle 23, the calculated success rate at high solar wind speed, although clearly above 50 %, was indicative rather than conclusive.

Estimated values of various statistical parameters used to evaluate the predictive model are presented in Table 2. As reported in Sect. 4, (PODy) values obtained during the rise/maximum/decay phases of the cycle and using the composite sample were 0.85, 0.64, 0.79 and 0.77, respectively, and these figures indicate the robustness of the modeling.

A representative reference metric defined by $[(\text{hits} + \text{correct nulls}) \times 100] / (\text{total number of predictions})$ is used to describe the success rates of the predictions relative to the measurements and application of a χ^2 test yields corresponding levels of statistical significance for the various sets/subsets investigated.

5.2 Possible influence on the predictions of the phase of the solar cycle

The data of Table 3 (columns 8–10) suggest that variations over a solar cycle in the environmental conditions attending shock transit (e.g. due to a change in the complexity of the interplanetary background at solar minimum and/or the changing helio-latitudes of flare initiation as a cycle progresses) might have been responsible for the improved performance shown by the model in the decay phase of solar cycle 23. See Fig. 1 which plots (following Table 3) the success rate *sr* (in green, left co-ordinate) and p_3 (in red, right co-ordinate) for the events measured during each of the three cycle phases. The above interpretation of the data should,

Table 10. 2 × 2 contingency table and associated loss table (denoted by c_{ij}), following Mozer and Briggs (2003).

		Observed	
		Yes (Y = 1)	No (Y = 0)
Forecasted	Yes (X = 1)	n_{11} (0)	n_{01} (c_{01})
	No (X = 0)	n_{10} (c_{10})	n_{00} (0)

however, be treated with caution since the decay phase was characterized by the availability of an increased number of sources of Type II radio data which could enhance the associated success of event detection. Also, a gradual gain in skill in using the model by the forecasters over a decade could be a contributory factor.

On the other hand, the consistently lower performance in the subsets of the model at solar maximum relative to the rise and decay phases (see the data displayed in the several tables in Sect. 4) may indicate that an inhibiting effect on the prediction outcome due to complex interplanetary conditions between the Sun and Earth is indeed real. It is noted that the decay phase of solar cycle 23 was characterized by an unusual period of reduced solar activity and low interplanetary magnetic field strengths (Schwadron et al., 2010). Also, these authors reported that the extended solar minimum of 2009–2010, which extended between solar cycles 23 and 24, was marked by fluxes of galactic cosmic ray radiation at near to their highest level in 25 years. It can be expected that the unusually quiet solar conditions pertaining during the decay phase of cycle 23 would favor the detection of solar shock events to a somewhat greater extent that would be the case during the rise phase of the cycle and this scenario is generally suggested by the statistics presented in the tables of Sect. 4.

5.3 The Briggs-Rupert skill score

Given that the consequence in financial terms of a hit or a miss depends on how a forecast is being used, Mozer and Briggs (2003) described a procedure that a user of shock forecasts can follow in order to calculate error estimates for the predictions that indicate the cost to that user of acting on an incorrect forecast. The methodology concerned was developed by M. Briggs and D. Rupert based on earlier work by Thomson (2000).

The Briggs-Rupert (BR) skill score which is based on the 2 × 2 contingency table shown in Table 10 is:

$$K_\theta = \frac{n_{11}(1 - \theta) - n_{01}\theta}{(n_{11} + n_{10})(1 - \theta)} \tag{8}$$

$$\theta = \frac{c_{01}}{c_{01} - c_{10}} \tag{9}$$

where c_{01} is the cost of a false negative forecast and c_{10} is the cost of a false positive forecast.

Since costs are user specific, different decision makers can potentially assign different values to these quantities and the utility of a forecast as determined using K_{θ} can in consequence be different for different users. For further details see Appendix A of Mozer and Briggs (2003).

5.4 Status of predictive modeling

The statistics presented in this paper provide information gathered over a solar cycle concerning the performance of HAFv.2 in predicting shock arrival times at Earth. These data comprise a detailed body of data of potential value to those engaged in the area of space weather operations, pending the availability of results that will ultimately be provided by those 3-D MHD models that start at the Sun and continue ‘seamlessly’ to the ionosphere and beyond (Sect. 1). These latter models, within which HAFv.2 will be subsumed, can be expected to provide in the long term greater insight into the physics underlying the propagation, evolution and interaction of solar wind disturbances on their way to Earth than is presently available. However, the difficulties of implementing such modeling (which requires the coupled simulation of phenomena that occur on vastly different spatial and temporal scales using highly synchronized codes), together with the problem as to how measured data can be assimilated into the models to allow their validation, presently pose major challenges which are not as yet resolved.

6 Conclusions

- Predictions of the arrivals at Earth of 584 shock events recorded at L1 over a full solar cycle (no. 23) have been estimated using the HAFv.2 model. The sample of shocks utilized was composed of 245 events recorded during the decay phase of the cycle, which were compared with previously available data from the rise (173 events) and maximum (166 events) phases.
- SATs were estimated using a ± 24 h hit window. For a random set of shock arrival times distributed within such a window a RMS error of 14.2 h is expected. The RMS errors yielded by HAFv.2 using data measured at L1 during the rise/maximum/decay phases of cycle 23 and on employing the composite sample were 11.6, 11.4, 10.6 and 11.4, respectively. These values were thus, in every case, better than would be expected for a random set.
- The complementary percentage success rates were 53 %, 51 %, 59 % and 55 %, respectively. Application of a χ^2 test yielded high levels of significance for model performance during the rise and minimum phases of the cycle as well as using the composite sample. At cycle maximum the significance level was marginal and this may reflect the influence of those complex conditions pertaining in interplanetary space during this period.
- The performance of the HAFv.2 model with respect to corresponding predictions of shock arrival times in association with a population of minor flares (197) showed a low level of significance during the various cycle phase. The figure of merit for the composite sample was $sr = 53\%$ with $p_3 = 0.1392$. The large number of correct nulls identified contributed to producing this result.
- The performance of HAFv.2 in association with predicting shocks characterized by short ($\tau \leq 20$ m) proxy piston driving times showed a low level of significance. There were substantially fewer events (33) with short driving times in the decay phase than during the rise (41) and maximum phases (54), which may reflect the nature of the flaring produced as the cycle evolved.
- Subsets formed to compare shocks that travelled in an interplanetary medium featuring solar wind speed (>400 km/s) and quiet flows (≤ 400 km s⁻¹) suggest a slightly enhanced success rate when the background solar wind was fast although the level of statistical significance at solar maximum was relatively low.
- The performance of HAFv.2 with respect to predicting shocks associated with flares located at heliolongitudes $>20^\circ$ and at heliolongitudes $\leq 20^\circ$ show that a substantially higher number of measured events originated at high than at low helio-longitudes (424 vs. 160) during solar cycle 23. A preponderance of events occurred in each case during the decay phase. Although in both instances there were a large number of false alarms, the success rate in this phase was 58 % at high solar longitudes and 62 % at low solar longitudes, each with a high level of statistical significance.
- The performance of HAFv.2 in predicting shocks with speeds <1200 km s⁻¹ at helio-longitudes $>20^\circ$ indicate that for 327 events in this category sr was 56 % and the result had a high level of significance. Due to the low number of events recorded with speeds ≥ 1200 km s⁻¹ the corresponding results were not statistically significant.
- The performance of HAFv.2 in predicting shocks with speeds <1200 km s⁻¹ originating at helio-longitudes $>80^\circ$ had an sr value of 60 % during the decay phase of the cycle but showed marginal statistical significance. The composite (89 events) sample was also characterized by an sr value of 60 % and it had a p_3 value of 0.1878.
- The statistical significance of the results obtained using the various sets/subsets was generally very low and

these results were not significant as compared with the hit by chance rate (50%), thus providing a low level of confidence in the predictions of the model with no compelling result encouraging its use. However, the data suggested that the success rates of HAFv.2 were higher when the background solar wind speed at the time of shock initiation was relatively fast. Therefore in scenarios where the background solar wind speed is elevated and the calculated success rate significantly exceeds the rate by chance, the forecasts could provide potential value to the customer. With the composite statistics available for solar cycle 23, the calculated success rate at high solar wind speed, although clearly above 50%, was indicative rather than conclusive.

- The parameter “Probability of Detection, yes” (POD_y) which presents the Proportion of Yes observations that were correctly forecast (i.e. the ratio between the shocks correctly predicted and all the shocks observed), yielded values for the rise/maximum/decay phases of the cycle and using the composite sample of 0.85, 0.64, 0.79 and 0.77, respectively.
- The consistently lower performance of the model at solar maximum relative to the rise and decay phases of the cycle in the various sets and subsets suggests an inhibiting effect on the prediction outcome due to the presence of complex interplanetary conditions between the Sun and Earth. Also, it can be expected that the unusually quiet solar conditions that pertained during the decay phase of solar cycle 23 would favor shock event detection to a somewhat greater extent that would be the case during the rise phase of the cycle. This scenario is also generally suggested by the statistics.
- The results obtained from detailed analysis of phase related sets and subsets constitute a resource for experimenters concerned with forecasting shock arrivals at Earth during solar cycle 24.
- Shock predictions are utilized in making commercially significant operational decisions (with regard, for instance, to placing a spacecraft into hibernation to protect it from an expected extreme solar event). Thus, use of the Briggs-Rupert skill score to estimate the loss incurred by a particular user through responding to an incorrect forecast can provide support in operational decision making.

Acknowledgements. KK wishes to acknowledge grant agency VEGA, project 2/0081/10 for support.

Guest Editor M. Gedalin thanks T. Nakagawa, A. Lui, and two other anonymous referees for their help in evaluating this paper.

References

- Akasofu, S.-I.: Predicting space weather storms as a space weather project, in: Geophysical Monograph Series, vol. 125, edited by: Song, P., Siscoe, G., and Singer, H., Publ. AGU, Washington D.C., pp. 320–337, 2001.
- Arge, C. N. and Pizzo, V. J.: Improvement in the prediction of solar wind conditions using near-real-time solar magnetic field updates, *J. Geophys. Res.*, 105, 10465–10479, 2000.
- Arge, C. N., Luhmann, J. G., Odstrcil, D., Schrijver, C. J., and Li, Y.: Stream structure and coronal sources of the solar wind during the 12th May 1997 CME, *J. Atmos. Solar Terr. Phys.*, 66, 1295–1309, 2004.
- Baker, D. N., Weigel, R. S., Rigler, E. J., McPherron, R. L., Vasiliadis, D., Arge, C. N., Siscoe, G. L., and Spence, H. E.: Sun-to-magnetosphere modeling: CISM forecast model development using linked empirical models, *J. Atmos. Solar Terr. Phys.*, 66, 1491–1497, 2004.
- Baker, D. N., Wiltberger, M. J., Weigel, R. S., and Elkington, S.: Present status and future challenges of modeling the Sun-Earth end-to-end system, *J. Atmos. Solar Terr. Phys.*, 69, 3–17, doi:10.1016/j.jastp.2006.07.017, 2007.
- Detman, T., Smith, Z., Dryer, M. Fry, C. D., Arge, N. C., and Pizzo, V.: A hybrid heliospheric modeling system: Background solar wind, *J. Geophys. Res.*, 111, A07102, doi:10.1029/2005JA011430, 2006.
- Dryer, M.: Interplanetary shock waves generated by solar flares, *Space Sci. Revs.*, 15, 403–468, 1974.
- Dryer, M. and Smart, D. F.: Dynamical models of coronal transients and interplanetary disturbances, *Adv. Space Res.*, 4, 291–301, 1984.
- Dryer, M., Fry, C. D., Sun, W., Deehr, C. S., Smith, Z., Akasofu, S.-I., and Andrews, M. D.: Prediction in real-time of the 2000 July 14 heliospheric shock wave and its companions during the “Bastille” epoch, *Sol. Phys.*, 204, 265–284, doi:10.1023/A:1014200719867, 2001.
- Dryer, M., Smith, Z., Fry, C. D., Sun, W., Deehr, C. S., and Akasofu, S.-I.: Real-time shock arrival predictions during the “Halloween 2003” epoch, *Space Weather*, 2, S09001, doi:10.1029/2004SW000087, 2004.
- Fry, C. D., Sun, W., Deehr, C. S., Dryer, M., Smith, Z., Akasofu, S.-I., Tokumaru, M., and Kojima, M.: Improvements to the HAF solar wind model for space weather predictions, *J. Geophys. Res.*, 106, 20985–21001, 2001.
- Fry, C. D., Dryer, M., Smith, Z., Sun, W., Deehr, C. S., and Akasofu, S.-I.: Forecasting solar wind structures and shock arrival times using an ensemble of models, *J. Geophys. Res.*, 108, 1070, doi:10.1029/2002JA009474, 2003.
- Fry, C. D., Dryer, M., Sun, W., Detman, T. R., Smith, Z., Deehr, C. S., Wu, C.-C., Akasofu, S.-I., and Berdichevsky, D.: Solar observation-based model for multi-day predictions of interplanetary shock and CME arrivals at Earth, *IEEE Trans. Plasma Sci.*, 32, 1489–1495, 2004.
- Fry, C. D., Dryer, M., Sun, W., Deehr, C. S., Smith, Z., Detman, T. R., Aran, A., Lario, D., Sanahuja, B., and Akasofu, S.-I.: Key links in space weather: Forecasting solar generated shocks and proton acceleration, *AIAA J.*, 43, 987–993, 2005.
- Fry, C. D., Detman, T. R., Dryer, M., Smith, Z., Sun, W., Deehr, C. S., Akasofu, S.-I., Wu, C.-C., and McKenna-Lawlor, S.: Real time solar wind forecasting: capabilities and challenges, *J. At-*

- mos. *Solar Terr. Phys.*, 69, 109–115, 2007.
- Hakamada, K. and Akasofu, S.-I.: Simulation of three-dimensional solar wind disturbances and resulting geomagnetic storms, *Space Science Revs.*, 31, 3–70, 1982.
- Lario, D., Sanahuja, B., and Heras, A. M.: Energetic particle events: Efficiency of interplanetary shocks at $50 \text{ keV} < E < 100 \text{ MeV}$ proton accelerators, *Astrophys. J.*, 509, 415–434, 1998.
- Lee, C. O., Luhmann, J. G., Odstrcil, D., MacNeice, P. J., Pater, I. de, Riley, P., and Arge, C. N.: The Solar Wind at 1 AU during the declining phase of solar cycle 23: Comparison of 3D numerical model results with observations, *Solar Phys.*, 254, 155–183, 2009.
- Luhmann, J. G.: CMEs and Space Weather, in: *Coronal Mass Ejections*, *Geophys. Monog. Ser.*, vol. 99, edited by: Crooker, N., Joselyn, J. A., and Feynman, J., Publ. AGU Washington D.C., pp. 291–299, 1997.
- Manchester IV, W. B., Vourlidas, A., Tóth, G., Lugaz, N., Roussev, I. I., Sokolov, I. V., Gombosi, T. I., De Zeeuw, D. L., and Opher, M.: Three dimensional MHD simulation of the 2003 October 28 coronal mass ejection: comparison with LASCO coronagraph observations, *Astrophys. J.*, 684, 1448, doi:10.1086/590231, 2008.
- McKenna-Lawlor, S. M. P., Dryer, M., Smith, Z., Kecskemety, K., Fry, C. D., Sun, W., Deehr, C. S., Berdichevsky, D., Kudela, K., and Zastenker, G.: Arrival times of Flare/Halo CME associated shocks at the Earth: comparison of the predictions of three numerical models with these observations, *Ann. Geophys.*, 20, 917–935, doi:10.5194/angeo-20-917-2002, 2002.
- McKenna-Lawlor, S. M. P., Dryer, M., Fry, C. D., Sun, W., Lario, D., Deehr, C. S., Sanahuja, B., Afonin, V. A., Verigin, M. I., and Kotova, G. A.: Predictions of energetic particle radiation in the close Martian environment, *J. Geophys. Res.*, 110, A03102, doi:10.1029/2004JA010587, 2005.
- McKenna-Lawlor, S. M. P., Dryer, M., Kartalev, M. D., Smith, Z., Fry, C. D., Sun, W., Deehr, C. S., Kecskemety, K., and Kudela, K.: Near real-time predictions of the arrival at Earth of flare-related shocks during Solar Cycle 23, *J. Geophys. Res.*, 111, A11103, doi:10.1029/2005JA011162, 2006.
- McKenna-Lawlor, S. M. P., Dryer, M., Fry, C. D., Smith, Z. K., Intriligator, D. S., Courtney, W. R., Deer, C. S., Sun, W., Kecskemety, K., Kudela, K., Balaz, J., Barabash, S., Futaana, Y., Yamauchi, M., and Lundin, R.: Predicting interplanetary shock arrivals at Earth, Mars and Venus: A real-time modeling experiment following the solar flares of 5–14 December, 2006, *J. Geophys. Res.*, 113, A06101, doi:10.1029/2007JA012577, 2008.
- Mozer, J. B. and Briggs, W. M.: Skill in real-time solar wind shock forecasts, *J. Geophys. Res.*, 108, 1262, doi:10.1029/2003JA009827, 2003.
- Odstrcil, D.: Modelling 3-D solar wind structure, *Adv. Space Res.*, 32, 497–508, 2003.
- Owens, M. J.: Combining remote and in situ observations of coronal mass ejections to better constrain magnetic cloud reconstruction, *J. Geophys. Res.*, 113, A12102, doi:10.1029/2008JA013589, 2008.
- Riley, P., Linker, J. A., and Mikic, Z.: An empirically-driven global MHD model of the solar corona and inner heliosphere, *J. Geophys. Res.*, 98, 15889–15901, 2001.
- Schaefer, J. T.: Critical success index as an indication of warning skill, *Weather and Forecasting*, 3, 570–575, 1990.
- Schwadron, N. A., Boyd, A. J., Kozarev, K., Golightly, M., Spence, H., Townsend, L. W., and Owens, M.: Galactic cosmic ray radiation hazard in the unusual extended solar minimum between solar cycles 23 and 24, *Space Weather*, 8, S00E04, doi:10.1029/2010SW000567, 2010.
- Smart, D. F. and Shea, M. A.: A simplified technique for estimating the arrival time of solar flare-initiated shocks, in: *Proc. STIP Workshop on solar/interplanetary intervals*, edited by: Shea, M. A., Smart, D. F., and McKenna-Lawlor, S., Publ. Book Crafters, Chelsea, Michigan, pp. 139–256, 1984.
- Smart, D. F. and Shea, M. A.: A simplified model for timing the arrival of solar flare-initiated shocks, *J. Geophys. Res.*, 90, 183–190, 1985.
- Smith, Z. and Dryer, M.: MHD study of temporal and spatial evolution of simulated interplanetary shocks in the ecliptic plane within 1 AU, *Sol. Phys.*, 129, 387–405, 1990.
- Smith, Z., Murtagh, W., Detman, T. R., Dryer, M., Fry, C. D., and Wu, C.-C.: Study of solar-based inputs into space weather models that predict interplanetary shock arrivals at Earth, *Proc. SCOSTEP Int. Solar Cycle Studies Symp.*, Tatranska, Lomnica, Slovakia, 23–28 June 2003, ESA SP-535, Publ. Noordwijk, The Netherlands, pp. 547–552, 2003.
- Smith, Z., Detman, T. R., Dryer, M., Fry, C. D., Wu, C.-C., Sun, W., and Deehr, C. S.: Verification method for space weather forecasting models using solar data to predict intervals of interplanetary shocks at Earth, *IEEE Plasma Phys. Trans. Plasma Sci.*, 32, Part I of III, 1498–1505, 2004.
- Smith, Z., Dryer, M., and Fry, C. D.: Determining shock velocities for inputs to Sun-to-Earth models from radio and coronagraph data, *Space Weather*, 3, S07002, doi:10.1029/2004SW000136, 2005.
- Smith, Z. K., Dryer, M., McKenna-Lawlor, S. M. P., Fry, C. D., Deehr, C. S., and Sun, W.: Operational Validation of HAFv2's Predictions of Interplanetary Shock Arrivals at Earth: Declining phase of Solar Cycle 23, *J. Geophys. Res.*, 114, A05106, doi:10.1029/2008JA013836, 2009.
- Sun, W., Akasofu, S.-I., Smith, Z. K., and Dryer, M.: Calibration of the kinematic method of studying the solar wind on the basis of a one-dimensional MHD solution and a simulation study of the heliosphere between November 22–December 6, 1977, *Planet. Space Sci.*, 33, 933–943, 1985.
- Sun, W., Dryer, M., Fry, C. D., Deehr, C. S., Smith, Z., Akasofu, S.-I., Kartalev, M. D., and Grigorov, K. G.: Evaluation of solar Type II radio burst estimates of initial solar wind shock speed using a kinematic model of the solar wind on the April 2001 solar wind swarm, *Geophys. Res. Lett.*, 29, 1171, doi:10.1029/2001GL013659, 2002a.
- Sun, W., Dryer, M., Fry, C. D., Deehr, C. S., Smith, Z., Akasofu, S.-I., Kartalev, M. D., and Grigorov, K. G.: Real-time forecasting of ICME shock arrivals at L1 during the “April Fool’s Day” epoch: 28 March–21 April 2001, *Ann. Geophys.*, 20, 937–945, doi:10.5194/angeo-20-937-2002, 2002b.
- Sun, W., Deehr, C. S., Fry, C. D., Dryer, M., Smith, Z., and Akasofu, S.-I.: Plane-of-sky simulations of interplanetary shock waves, *Geophys. Res. Lett.*, 30, 2044, doi:10.1029/2003GL017574, 2003.
- Thomson, A. W. P.: Evaluating space weather forecasts of geomagnetic activity from a user perspective, *Geophys. Res. Lett.*, 27, 4049–4052, 2000.

- Tóth, G., Sokolov, I. V., Gombosi, T. I., Chesney, D. R., Clauer, C. R., De Zeeuw, D. L., Hansen, K. C., Kane, K. J., Manchester, W. B., Oehmke, R. C., Powell, K. G., Ridley, A. J., Rousev, I. I., Stout, Q. F., Volberg, O., Wolf, R. A., Sazykin, S., Chan, A., Yu, Bin, and Kóta, J.: Space Weather Modeling Framework: A new tool for the space science community, *J. Geophys. Res.*, 110, A12226, doi:10.1029/2005JA011126, 2005.
- Tóth, G., De Zeeuw, D. L., Gombosi, T. I., Manchester IV, W. B., Ridely, A. J., Sokolov, I. V., and Rousev, I. I.: Sun-to-thermosphere simulation of the 28-30 October 2003 storm with the Space Weather Modeling Framework, *Space Weather*, 5, S06003, doi:10.1029/2006SW000272, 2007.
- Wang, Y. M. and Sheely, N. R.: Solar wind speed and coronal flux tube expansion, *Astrophys. J.*, 355, 726–732, 1990.