course be changed if the sample is clearly defined to be a “training set”. During the course of this study, both HAFv.1 and HAFv.2 were used in real time at different times as part of the three-model ensemble. Thus, it was decided, in view of the model differences between the two versions, to consider the N = 173 sample as a “training set” in real time and to study the two HAFv.2 thresholds afterward in order to tune the Shock Searching Index. SSI (log of the normalized momentum flux change at each time step) and the absolute value of the momentum flux change; subflares and 1F flares were also deleted. Also, the STOA and ISPM models were examined retrospectively in order to eliminate interacting events, reducing their sample size to N = 148. The HAFv.2 sample remained at N = 173 because it can consider interacting events.

Thus, the Contingency Table 2, compiled from the real time experience, is as follows for the three model ensemble (where HAFv.1 was used initially and switched later to HAFv.2):

Table 2. Real-Time Contingency Tables

<table>
<thead>
<tr>
<th>STOA</th>
<th>Obs</th>
<th>ISPM</th>
<th>Obs</th>
<th>HAF</th>
<th>Obs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pred</td>
<td>Y 53 57</td>
<td>Pred</td>
<td>Y 39 31</td>
<td>Pred</td>
<td>Y 50 78</td>
</tr>
<tr>
<td>N 15 48</td>
<td>N 29 74</td>
<td>N 18 27</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Clearly, the HAFv.1 and HAFv.2 models performed poorly in comparison to the other two models during their “training” exercise. The retrospective study, as outlined above, is given in Table 3 below.

Table 3. Retrospective Contingency Tables

<table>
<thead>
<tr>
<th>STOA</th>
<th>Obs</th>
<th>ISPM</th>
<th>Obs</th>
<th>HAF</th>
<th>Obs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pred</td>
<td>Y 49 44</td>
<td>Pred</td>
<td>Y 36 24</td>
<td>Pred</td>
<td>Y 48 41</td>
</tr>
<tr>
<td>N 12 43</td>
<td>N 25 63</td>
<td>N 20 64</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The “tuning” of HAFv.2 included a threshold of SSI ≥ -0.35 and elimination of optical subflares and 1F classified flares in soft X-rays except for the 1F flare shock interactions with CIR as they impact the spacecraft. The RMS errors of all models (~12 hr) are almost identical. This result agrees with the theoretical results of Heinemann (2001), based on velocity and density uncertainties in the upstream solar wind, and suggests that further improvement may not be expected by these or any other model in the shock arrival time category. Additional skill study indicates that the probability of a “Yes” detection (PODy) of shock arrival is 80% for STOA and 59% for ISPM when the above Table 3 is corrected for non-interacting events. Whereas, for HAFv.2, it is 71% for all events. Alternatively, the probability of a “No” detection (PODn) is, respectively, 49%, 72% and 61% for this ensemble. Also, the false alarm rate is, again respectively, 47%, 40% and 46%; and the Heidke Skill Score (that corrects for the probability of detecting a shock by chance) is 28%, 32% and 30%. While the three models appear to be nearly equal in some respects, the HAFv.2’s advantage is in the provision of a global view of multiple, interacting events. HAFv.2 is superior in its ability to sort out shock interactions and CIRs, but it is certainly no better than the other two in calculating SATs in an inhomogeneous solar wind. However the “tuned” version has yet to be tested in the real time mode.

7. CONCLUDING REMARKS

This review took the perspective of a modeller who is concerned with the initialization of solar events as well as the necessity to incorporate realistic geometries (e.g., 3D MHD). The former concern remains with us as we strive to use available real-time solar observations to help with this quest. Comments were made on the return of a wider set of initializing possibilities with the apparent demise of the “solar flare myth” that temporarily inhibited these choices. The 3D MHD simulation of at least one case, the well-observed EIT wave, followed by a halo CME was mentioned as a case in point. Thus the mimicked “flare” is now considered as a very realistic possibility. Some historical 1D MHD initializations during the August 1972 flare events were mentioned together with the contemporary use, again outside the critical solar wind locations, of the Bastille ICME’s evolution out to 63 AU from July to November 2000. Classical similarity theory was also invoked as a benchmark to test the ‘energy–shock travel time’ studies via numerical methods. A kinematic model was also introduced via its use in real-time forecasting of the Sun-to-Earth shock arrival times as well as to illustrate the multiple interactions of flare-generated shocks during April 2001 and the pre-existing, non-uniform, stream-stream interactions. Finally, we discussed the real-time (and retrospective) contingency tables of three models (STOA, ISPM and HAFv.2) presently being used, again in real-time, in shock arrival forecasting.

We have recommended upgrading of the HAFv.2 model, firstly, via a hybrid approach that includes a 3D MHD code and, secondly, by its eventual replacement of a 3D MHD code (still with a single, ideal, infinitely-conducting plasma) that starts at the Sun. This plan (c.f., Dryer, 1998) should incorporate appropriate, time-dependent, boundary conditions that take projected characteristics into account within subsonic
and sub-Alfvénic flows. Thus, the ultimate objective of accurate Bz forecasting, as well as of the other physical and derived parameters, can be achieved in a timely and orderly fashion.

ACKNOWLEDGMENT

The writer wishes to thank the following scientists for their gracious help and assistance without which this paper would not have been completed: Ghee Fry, Zdenka Smith, Charles Deehr, Sun Wei, Syun Akasofu, Bill Ostdrirch, Tom Detman and Eduardo Araujo-Pradera. MD is also grateful for the hospitality of the NOAA Space Environment Center.

8. REFERENCES


Smith, Z., M. Dryer, E. Ort and W. Murtagh, Performance of interplanetary shock prediction models: