Ensemble Forecasting of Major Solar Flares

J.A. Guerra*, A. Pulkkinen, & V.M. Uritsky

The Catholic University of America and NASA Goddard Space Flight Center

ESWW 12. Session - Solar Storms: Flares, CMEs and Solar Energetic Particle (SPE) Events

*Now at Trinity College Dublin.
Motivation

• ASSA, ASAP, and MAG3 are hosted at NASA CCMC.
• CCMC forecasters also have NOAA forecasts.

Four (sometimes different) forecasts for the same photospheric conditions, Can they be combined?
Purpose

1. Demonstrate that an ensemble forecast can be constructed and perform better than any ensemble member.
2. Find the conditions (internal parameters) for the ensemble forecast to perform better than their members.
3. Demonstrate the applicability to real-time prototyping environment.
Ensemble Forecast

Ensemble of methods:

\{MAG4, ASSA, ASAP, NOAA\}

Linear combination of probabilities

\[ P^c(w, t) = w_{MAG4}P_{MAG4}(t) + w_{ASSA}P_{ASSA}(t) + w_{ASAP}P_{ASAP}(t) + w_{NOAA}P_{NOAA}(t) \]

with \( w_{MAG4} + w_{ASSA} + w_{ASAP} + w_{NOAA} = 1 \).

Ensemble models: Determining \( w \)'s

1. Average over active region sample (AVE)
2. Analysis of Extended Time Series (ETS)
3. Equal weights
4. Human-influenced vs Automated
Time Series Constructions

Probability (solid lines): at each $t$, probability for a 24-h prediction window. Events (dotted lines): a flare observed between $t$ and $t + 24$-h.
Metric Optimization

To calculate \( w \)'s: 1) Linear Combination of P's, 2) Apply threshold \( (P_{th}) \), 3) Calculate HSS, 4) Repeat until HSS is maximized.

\[
HSS = \frac{a + d - e}{n - e}.
\]

2x2 Contingency table

<table>
<thead>
<tr>
<th>Event Forecast</th>
<th>Event</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>( a )</td>
<td>( b )</td>
</tr>
<tr>
<td>No</td>
<td>( c )</td>
<td>( d )</td>
</tr>
</tbody>
</table>

\( P_{th} = 25\% \)
Monte-Carlo Simulation

The metric Optimization process was inserted inside a MC type algorithm to calculate the weights.

Repeat N times

- ARs sample
- Random selection
- Training subsample
- Validation subsample

Optimization of HSS:
1. LC + \( th \)
2. 2x2 table
3. Calculate HSS

\( \{w^{max}\}_i \)

1. LC using \( \{w^{max}\}_i + th \)
2. 2x2 table
3. Calculate HSS and other metrics

AR sample: 13 ARs, 8 X-class, 64 M-class.
Results: Probabilistic Forecast

Probabilistic Forecast Attributes

<table>
<thead>
<tr>
<th>Forecasting method</th>
<th>( \langle P_i \rangle )</th>
<th>QR</th>
<th>REL</th>
<th>RES</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAG4</td>
<td>0.286</td>
<td>0.219</td>
<td>0.0487</td>
<td>0.0329</td>
<td>-0.0756</td>
</tr>
<tr>
<td>ASSA</td>
<td>0.261</td>
<td>0.194</td>
<td>0.0311</td>
<td>0.0397</td>
<td>0.0464</td>
</tr>
<tr>
<td>ASAP</td>
<td>0.197</td>
<td>0.212</td>
<td>0.0272</td>
<td>0.0193</td>
<td>-0.0398</td>
</tr>
<tr>
<td>NOAA</td>
<td>0.295</td>
<td>0.186</td>
<td>0.0145</td>
<td>0.0363</td>
<td>0.0882</td>
</tr>
<tr>
<td>Ensemble I</td>
<td>0.318</td>
<td>0.182</td>
<td>0.0092</td>
<td>0.0307</td>
<td>0.1066</td>
</tr>
<tr>
<td>Ensemble II</td>
<td>0.319</td>
<td>0.184</td>
<td>0.0109</td>
<td>0.0299</td>
<td>0.1127</td>
</tr>
<tr>
<td>Ensemble Equal</td>
<td>0.321</td>
<td>0.181</td>
<td>0.0076</td>
<td>0.0303</td>
<td>0.1103</td>
</tr>
</tbody>
</table>

\( M \)-class (Total events: 888, Climatology=0.285, \( \overline{QR} = 0.204 \))

\( X \)-class (Total events: 169, Climatology=0.054, \( \overline{QR} = 0.051 \))

\[
QR = \frac{1}{N_p} \sum_{i=1}^{N_p} (P_i - E_i)^2,
\]

\[
REL = \frac{1}{N_p} \sum_{i=1}^{T} N_{R_i} (\langle P_i \rangle - \langle E_i \rangle)^2,
\]

\[
RES = \frac{1}{N_p} \sum_{i=1}^{T} N_{R_i} (\langle E_i \rangle - \overline{E})^2
\]

Skill Score

\[
SS = \frac{\overline{QR} - QR}{\overline{QR}}
\]

For \( M \)-class flares all ensembles improved most attributes.
Results: Probabilistic Forecast

Reliability diagram for Equal weights ensemble

Ensemble forecast (left) for M-class flares follows the perfect reliability curve (diagonal line).
Results: Categorical Forecast

HSS as function of probability threshold ($P_{th}$).

We found:
- AVE ensemble similar to Equal ensemble.
- Best ensemble HSS: for 25% (M) and 15% (X)
- Highest M-class HSS for NOAA. Distribution of prob. forecasts is spread towards higher values.
Results: Categorical Forecast

Comparing HSS values for ensemble with only Automated (3M), all methods (4M), and Human-influenced (NOAA) forecasts.

We found:
- Inclusion of human-influenced forecasts improve the ensemble prediction in some range of $P_{th}$ (25 – 40 %)
Real Time Implementation

Prototype is currently under testing at CCMC.

---

<table>
<thead>
<tr>
<th>ProbabilityWindowStartTime</th>
<th>PredictionWindowDuration</th>
<th>FlareClass</th>
<th>AR</th>
<th>Probability</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015-04-03T17:00:00Z</td>
<td>P24H</td>
<td>M</td>
<td>12317</td>
<td>MAG4</td>
<td>1.</td>
</tr>
<tr>
<td>2015-04-03T17:00:00Z</td>
<td>P24H</td>
<td>M</td>
<td>12317</td>
<td>ASSA</td>
<td>4.</td>
</tr>
<tr>
<td>2015-04-03T17:00:00Z</td>
<td>P24H</td>
<td>M</td>
<td>12317</td>
<td>ASAP</td>
<td>0.</td>
</tr>
<tr>
<td>2015-04-03T17:00:00Z</td>
<td>P24H</td>
<td>M</td>
<td>12317</td>
<td>NOAA</td>
<td>5.</td>
</tr>
<tr>
<td>2015-04-03T17:00:00Z</td>
<td>P24H</td>
<td>X</td>
<td>12317</td>
<td>Ensemble1</td>
<td>2.</td>
</tr>
<tr>
<td>2015-04-03T17:00:00Z</td>
<td>P24H</td>
<td>X</td>
<td>12317</td>
<td>Ensemble2</td>
<td>2.</td>
</tr>
<tr>
<td>2015-04-03T17:00:00Z</td>
<td>P24H</td>
<td>X</td>
<td>12317</td>
<td>MAG4</td>
<td>0.</td>
</tr>
<tr>
<td>2015-04-03T17:00:00Z</td>
<td>P24H</td>
<td>X</td>
<td>12317</td>
<td>ASSA</td>
<td>0.</td>
</tr>
<tr>
<td>2015-04-03T17:00:00Z</td>
<td>P24H</td>
<td>X</td>
<td>12317</td>
<td>ASAP</td>
<td>0.</td>
</tr>
<tr>
<td>2015-04-03T17:00:00Z</td>
<td>P24H</td>
<td>X</td>
<td>12317</td>
<td>NOAA</td>
<td>1.</td>
</tr>
<tr>
<td>2015-04-03T17:00:00Z</td>
<td>P24H</td>
<td>X</td>
<td>12317</td>
<td>Ensemble1</td>
<td>0.</td>
</tr>
<tr>
<td>2015-04-03T17:00:00Z</td>
<td>P24H</td>
<td>X</td>
<td>12317</td>
<td>Ensemble2</td>
<td>0.</td>
</tr>
</tbody>
</table>

Example text output.
Concluding Remarks:

- Linear combination of predictions and track records of methods result in improvement of probabilistic and categorical flare forecasts.
- Human expertise and knowledge is still of great value in flare forecasting.
- Size and content of the forecasts/events sample affects performance.

More details: [http://dx.doi.org/10.1002/2015SW001195](http://dx.doi.org/10.1002/2015SW001195)