Metrics report on software for prediction of daily-averaged MeV electron intensity at geostationary orbit

To evaluate the software predicting daily-averaged MeV electron intensity at the geostationary orbit, we ran it for a period of of 84 days, from July 17 to October 7 (see table below), giving the model 27 days of lead time, using as input solar wind data from NOAA SEC ftp site. We calculated skill scores, comparing model output with three different reference models (see table below).

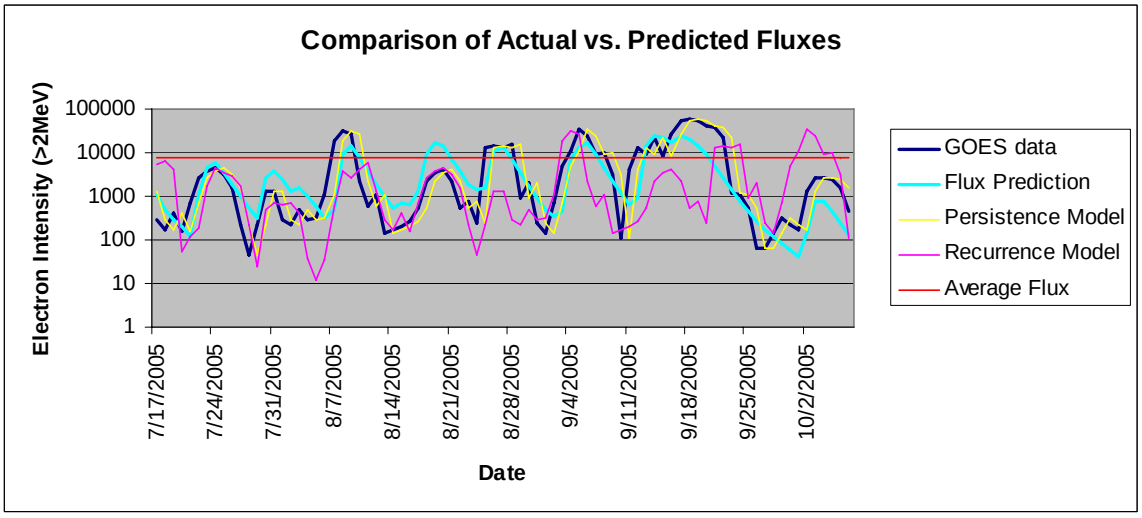
Predictions of MeV electron intensity (in the > 2.0 MeV range) for the next day produced by the model were evaluated using GOES 10 and 12 measurements downloaded from NOAA SEC ftp site. GOES 10 and 12 data were combined to form daily averaged fluxes.

Model Score:
D t = SQRT(Sum((H model - H data)**2) / N points)
D r = SQRT(Sum((H reference model - H data)**2) / N points)

Skill Score:
M = 1 - D t / D r

Reference Model	Skill Score
1-day persistence	-0.3544
Average	0.2041
27-day recurrence	0.2935

The graph below shows fluxes predicted by the software, GOES data and output from reference models for the period of July 17 – October 7.



Skill score comparison for quiet vs stormy period

We then compared skill scores during a quiet and a stormy period and looked at the effect of a day's worth of missing data on the prediction results.

To distinguish stormy periods from quiet ones, we used the geomagnetic Kp index recorded in the Space Environment Center (SEC) space weather alert archives, http://www.sec.noaa.gov/alerts/archive.html (see Fig. 1 below). When Kp index was 5 or higher we considered this an indication of storm.

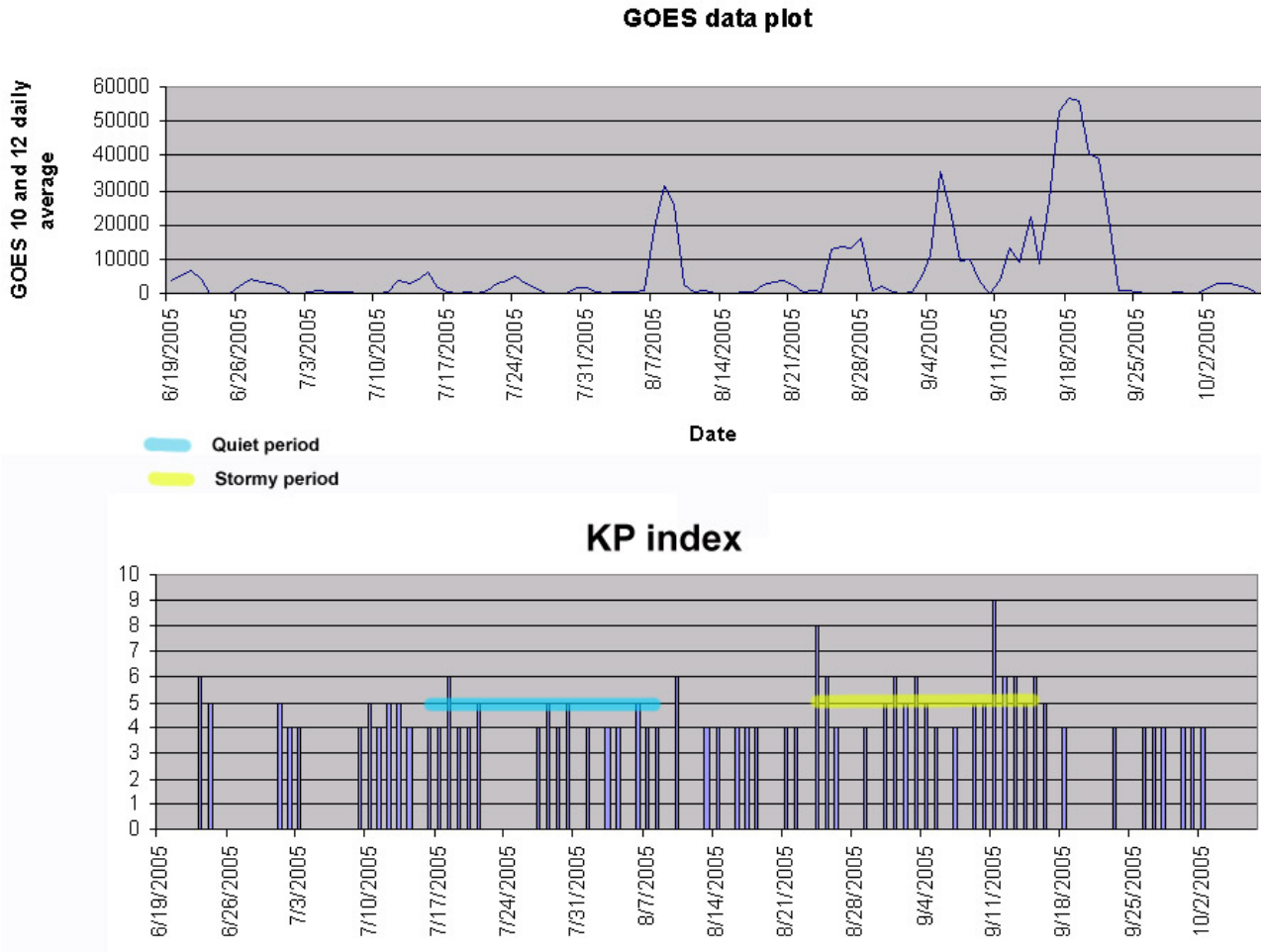


Figure 1

In our initial interval of calculations, July 17 to October 7, we selected two periods of 23 days each: what we considered <u>a quiet period</u> (July 17 to August 9), for wich the total sum of daily Kp values greater than 5 was equal to 26; and <u>a stormy period</u> (August 24 to September 16), for which the total sum of Kp >=5 was 88.

Reference Model	Skill Score		
	July 17-October 7	**Quiet:** July 17 to August 9	**Stormy:** August 24 - September 16
1-day persistence	-0.3544	-0.3864	0.2355
Average	0.2041	0.2294	0.2047
27-day recurrence	0.2935	0.2468	0.3629

Skills scores demonstrate that during a stormy period the model being evaluated produced better results compared to all reference models, while during quieter period 1-day persistence model gave a better result than the model being evaluated.

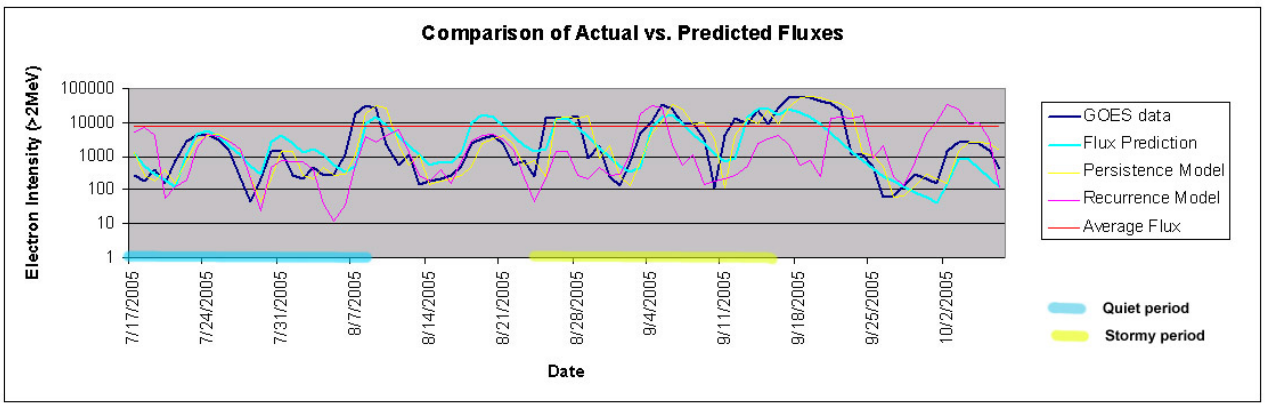


Figure 2. Actual and predicted fluxes during the quiet and stormy period

Effect of missing data on flux prediction

To see the effect of missing data for one day on the model prediction of electron flux, we introduced a day gap by leaving July 19 out of the list of days we specify for running in non-real time mode and compared this run with the full run. As a result, there was a change in predicted fluxes, however the effect of this gap can be seen only for two days following July 19th (see Figs. 3 and 4)

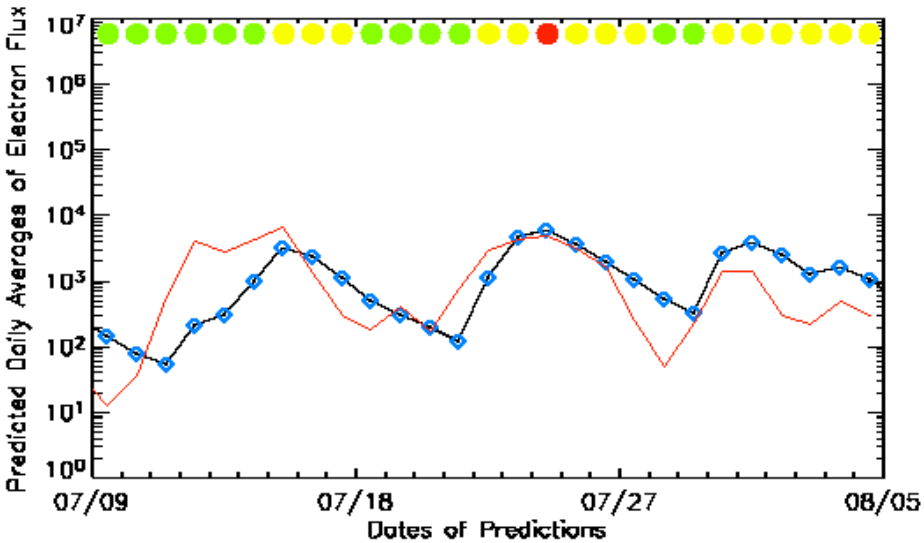


Figure 3: Predicted vs actual fluxes; no missing data.
(red is the actual GOES data; blue – model prediction)

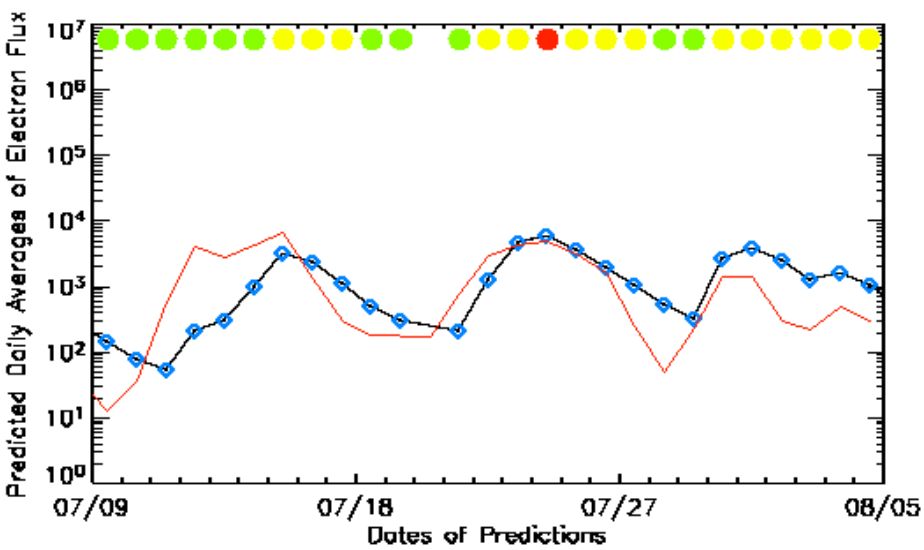


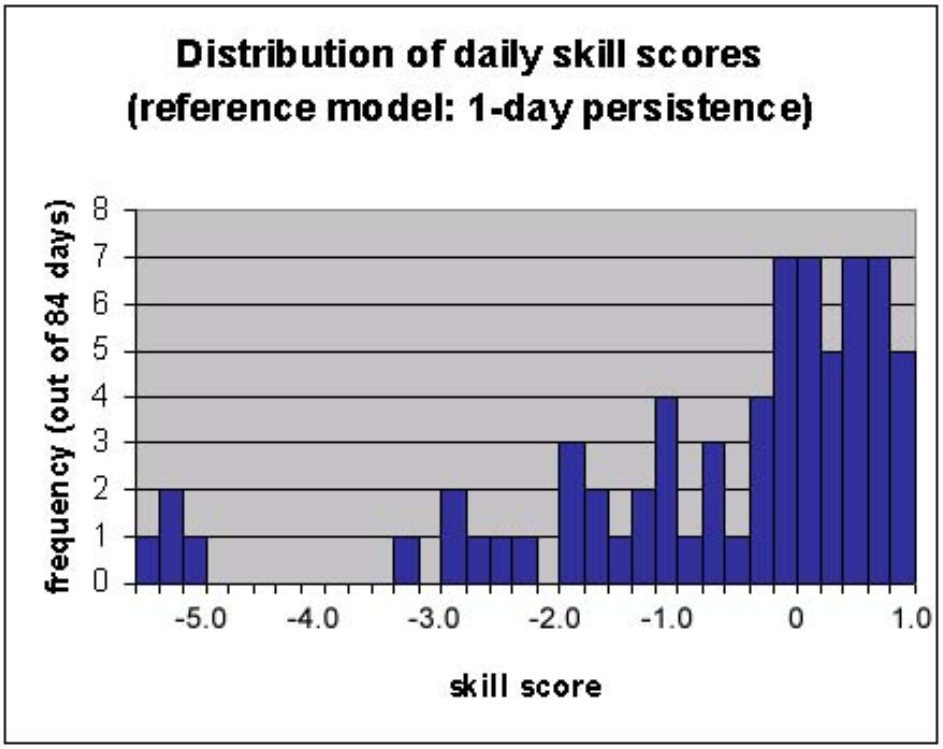
Figure 4: Predicted vs actual fluxes; July 19th data missing.
(red is the actual GOES data; blue – model prediction)

As the result, the effect of a day's worth of missing data on the skill score for the period of 84 days was negligible:

Reference Model	**Skill Score**	
	July 17-October 7	**Same time interval, but data missing for July 19**
1-day persistence	-0.3544	-0.3544
Average	0.2041	0.2026
27-day recurrence	0.2935	0.2971

Distribution of values for daily skill scores

Finally, we took a look at the distribution of values for daily skill scores (Fig. 1 below). The reference model was 1-day persistence. Scores above 0 indicate that the software being evaluated performed better than 1-day persistence reference model; negative scores indicated that persistence outperformed the UPOS software. In addition to skill scores shown, out of 84 days the scores for 13 days were below -6 and are not shown on the histogram below.



Model Score:
D t = (H model - H data) / N points
D r = (H reference model - H data) / N points

Daily skill Score:
M = 1 - D t / D r

The scattering of skill score and the error can have three sources:
- errors in input solar wind data
- errors in GOES electron flux measurements
- errors created by the arbitrary selection of the time interval