３.国際ワークショップ 目次詳細
International Workshop

（所属・職名は平成30年3月現在）
(Affiliation and Department are correct as of March 2018)

Operational Flare Forecasts: a systematic community comparison

Dr. KD Leka (Nagoya University / ISEE; NWRA/Boulder)

Purpose:

Solar flares are energetic and sudden releases of energy associated with complex magnetic fields in the lower solar atmosphere, with resulting near-Earth disturbances occurring within minutes of initial detection. True forecasting is required in the context of potential impacts to communication, positioning, and timing services. Today's operational algorithms generally provide probabilistic forecasts for flares above a defined Soft X-Ray threshold to occur within a given time interval. Routine operational forecasts have been produced for decades by (for example) the Space Weather Prediction Center of the US National Oceanic and Atmospheric Administration and the UK MetOffice. With the recognition of space weather as an international threat, new flare forecasting systems are being developed around the world. The purpose of this workshop was to directly compare the performance of these algorithms over a variety of agreed-upon operational settings (different event thresholds, validity periods, forecast latencies). The benchmarks developed would be then available for evaluating future methods.

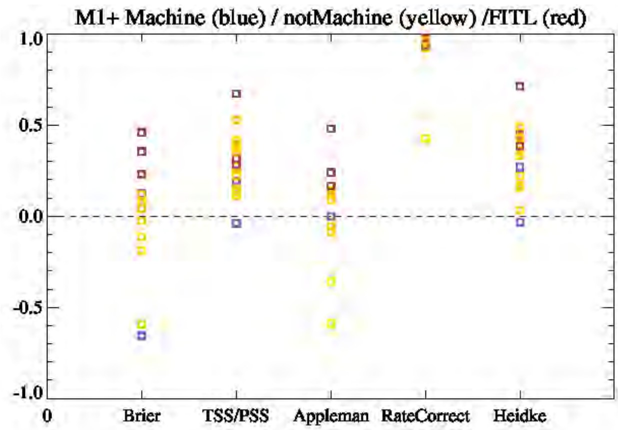


Figure 1: an example preliminary result from the "General Characteristics" analysis, showing the anonymous performance for numerous evaluation metrics as a function of whether Machine Learning algorithms were used, not used, or if there was a "Forecaster In The Loop." Surprisingly, the ML-based results were not routinely best.

Workshop Period / Place / Participants:

Participation was limited to only fully operationally-deployed systems, excluding those algorithms in a developmental or research stage. This still brought together representatives from 14 different qualifying methods. The face-to-face meeting was held 31 October – 02 November 2017 at ISEE / Nagoya University; some representatives participated remotely. Of note, all methods submitted forecasts ahead of time, with which Drs. Leka and Park prepared some preliminary results to enable fruitful discussion.

Results:

Topics discussed during the workshop were numerous and very technical in nature. “What constitutes an operational forecast” began the discussion; by consensus there is no requirement regarding human participation, but it was agreed that operational systems must provide forecasts “no matter what.” As such, missing forecasts in the submissions are flagged for explicit handling in the calculation of evaluation metrics. Other topics included combining categorical forecasts into exceedance forecasts, the impacts of variable climatology estimates (solar cycle variations), and accommodating differing forecast issuance times. As expected, the participants had a long discussion regarding metrics; it was decided that generally speaking the metrics which would be used would be those that did not require or impose probability thresholds, since that can unfairly penalize those methods that optimized using different thresholds. Hence, metrics would focus on evaluating the methods as probabilistic forecasts (using, e.g., the Receiver Operating Characteristic [ROC] plots, Reliability plots, Brier scores, Gini coefficients, and Ranked Probability Score [RPS]).

The initial evaluation results were initially presented in a completely anonymous manner in order to focus on broad performance characteristics rather than ranking or competition. Reliability plots, ROC curves, and standard skill scores were presented. Of note was the *wide* variation of performance, but also the generally poor performances overall. Forecast methods were grouped according to broad algorithm methodologies to investigate the causes of performance differences. One example from such a grouping is shown in Figure 1, which compares the performance of algorithms using Machine Learning techniques to those using other statistical classifiers or associations, and to those for which a human is still involved ("Forecaster In The Loop", or FITL). Of note, very few machine-learning based methods have transitioned to being fully operational. Separately, the large fraction of currently-running operational methods that are based on the historical McIntosh classification system for active regions contrasted with the small number of methods which perform independent quantitative evaluation of active region evolutionary, magnetic field, or model-corona characteristics.

Three publications are thus far planned (see Form 3-1): an analysis of performance results based on the numerous broad categories discussed, a short paper that indeed presents the methods' rankings according to numerous metrics (where it is almost guaranteed that the top performer by one metric will not be the top performer as judged by another metric), and an analysis of a recent "case study" of operational forecasts – with an in-depth analysis of the insights and difficulties associated with interpreting case studies. Numerous presentations of the results will be given at multiple international meetings in the upcoming months.

The workshop website, with data and analysis code, will eventually be released to the community. It will include the data needed for testing new evaluation metrics, and code to apply evaluation metrics to new method results in the same manner as done for the workshop-related publications. Thus, this workshop will provide a legacy of benchmarks for evaluating the performance of future operational flare-forecasting algorithms.

The Solar Cycle Prediction Workshop in Nagoya

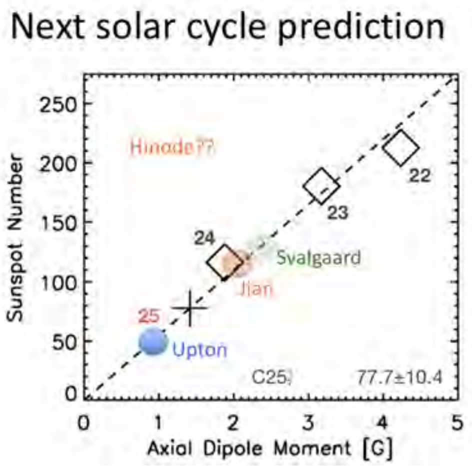
Shinsuke Imada (Nagoya University, ISEE)

It is thought that the longer-term variations of the solar activity may affect the Earth's climate. Therefore, predicting the next solar cycle is crucial for the forecast of the "solar-terrestrial environment". To build prediction schemes for the next solar cycle is a key for the long-term space weather study. Recently, the relationship between polar magnetic field at the solar minimum and next solar activity is intensively discussed. Because we can determine the polar magnetic field at the solar minimum roughly 3 years before the next solar maximum, we may discuss the next solar cycle 3years before. Further, the longer term (~5 years) prediction might be achieved by estimating the polar magnetic field with the Surface Flux Transport (SFT) model. Recently, we are developing a prediction scheme by SFT model and adapting to the Cycle 25 prediction. The predicted polar field strength of Cycle 24/25 minimum is several tens of percent smaller than Cycle 23/24 minimum. The result suggests that the amplitude of Cycle 25 is weaker than the current cycle. To evaluate the robustness of our result, it is very fruitful to compare the results done by the other foreign groups which use SFT model to predict. Below the list of members of the workshop who study next solar cycle prediction extensively.

Members:

L. Upton (HAO/NCAR)
R. Cameron (MaxPlanck Institute)
M. Dikpati (HAO/NCAR)
A. Munoz-Jaramilo (South West Research Institute)
L. Svalgaard (Stanford Univ.)
J. Jiang (Beihang Univ.)
D. Shiota (NICT)
H. Hotta (Chiba Univ.)
K. Kusano (Nagoya Univ.)
H. Iijima (Nagoya Univ.)
M. Fujiyama (Nagoya Univ.)
S. Imada (Nagoya Univ.)

The workshop was held from November 27th (Mon) to December 2nd (Saturday) at Nagoya University (Higashiyama Campus) Research Institute Building II Room 409. First two days, we have review each own research, and found several differences in the results. Below is the summary for the final prediction results in several groups.



There are three important detail pointss for comparison; 1) assumption used in the model, 2) numerical scheme, 3) initial and boundary conditions. The assumptions used in the model are different in some points. However, after several discussions, we conclude that the differences in assumption should not be the main reason for the differences in each result. Second, we have found that the numerical schemes which are used in the studies are all different. The effect itself is also very minor, we believe. To confirm it, we will test the cycle prediction problem by using the same initial and boundary conditions. Then, this point will be clarified. The last point is critical, we think. We will start to predict the next solar cycle by using the same initial and boundary conditions, and then discuss which condition affect the result deeply. These comparison study will be submitted to PASJ as a collaboration study paper.