Planetary K-index (Kp) Forecast Verification - Preliminary Results

N. H. Leftridge[1], R. A. Steenburgh[1,2]
University of Colorado, Boulder, CO
NOAA Space Weather Prediction Center, Boulder, CO

1. Abstract

The planetary K-index (Kp) is a measure of the state of the geomagnetic field. It is constructed every three hours using data gathered from 13 ground-based magnetic observatories. The National Oceanic and Atmospheric Administration's Space Weather Prediction Center (SWPC) in Boulder, Colorado, provides 0-72 hours forecasts of Kp twice a day at 00 and 12 UTC. In this study, we verify SWPCs 00 UTC, 0-24 hour forecasts using a variety of standard measures. In particular, we explore verification statistics for the 3 hourly values, the daily maximum Kp value, and the daily, binary storm, no-storm forecasts (Kp>=5). We found a broad tendency to over-forecast the 3-hourly Kp indices. The daily maximum forecasts and storm, no-storm forecasts showed a similar tendency.

2. Objectives

- Calculate forecast verification statistics for 24-hour forecasts of 3-hourly Kp, daily maximum Kp, and storms (Kp >=5)
- Examine the implications of using Kp forecasts to drive extended numerical space weather model predictions
- Identify any performance trends revealed by the verification data to inform forecast refinement, training needs
- Use results as a starting-point for discussions with sectors that use Kp forecasts

```
Product: 3-Day Forecast
Issued: 2017 Sep 15 0030 UTC
Prepared by the U.S. Dept. of Commerce, NOAA, Space Weather
Prediction Center

A. NOAA Geomagnetic Activity Observation and Forecast

The greatest observed 3 hr Kp over the past 24 hours was 6
(NOAA Scale G2).
The greatest expected 3 hr Kp for Sep 15-Sep 17 2017 is 6
(NOAA Scale G2).

NOAA Kp index breakdown Sep 15-Sep 17 2017

             Sep 15      Sep 16      Sep 17
00-03UT        6 (G2)        4           4
03-06UT        4             4           3
06-09UT        4             3           3
09-12UT        3             3           3
12-15UT        3             3           2
15-18UT        4             3           2
18-21UT        4             3           3
21-00UT        5 (G1)        4           3

Rationale: The geomagnetic field is expected to see an
isolated period of G2 (Moderate) geomagnetic storming on day
one (15 Sep) as CH HSS influences continue. Day two (16 Sep)
is expected to be below the G1 (Minor) storm threshold, but
could see a chance for an isolated period
of G1 storming early, due to continuing CH HSS effects.
```

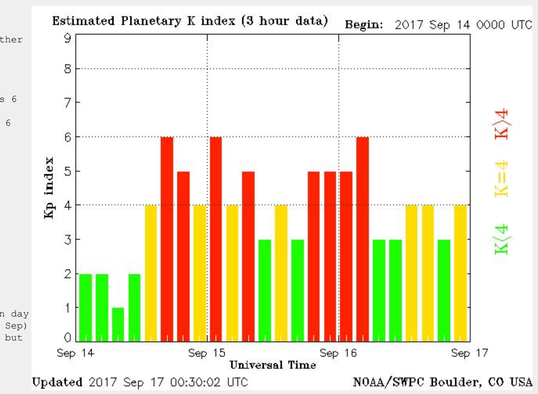

Figure 1: An example forecast and observation comparison. NOAA Space Weather Prediction Center 3-day forecast product (left) for 2017-09-15 issued at 0030 UT and corresponding Planetary K-index (Kp) plot centered on 15 September.

3. Background

Geomagnetic storms can interfere with technology on earth, the most susceptible include power and communication systems. These storms can be identified, quantified, and forecasted using the planetary K-Index (Kp). It is a quasi-logarithmic scale produced every 3 hours and describes global geomagnetic activity using readings from 13 magnetometers located between 44^o and 60^o North and South latitude. The index ranges from 0-9, and is broken into thirds (-,o,+) such that 9o indicates the highest level of geomagnetic activity, and 0o indicate the lowest level of geomagnetic activity.

In November, 2012, SWPC began producing a 3-day forecast product which included 3-hourly Kp forecasts out to 72hrs. To date, there have been limited attempts to verify the SWPC Kp forecast performance. SWPC is considering driving numerical space weather prediction models (e.g. WAM-IPE, CTIPe, Ovation with these forecast values, prompting this work.

Kp	0o	0+	1-	1o	1+	2-	2o	2+	3-	3o	3+	4-	4o	4+
ap	0	2	3	4	5	6	7	9	12	15	18	22	27	32
Kp	5-	5o	5+	6-	6o	6+	7-	7o	7+	8-	8o	8+	9-	9o
ap	39	48	56	67	80	94	111	132	154	179	207	236	300	400

Table 1: Table from the German Research Centre for Geosciences (GFZ) website showing the range of Kp values and the corresponding ap (linear scale) indices.

4. Method

- Compiled Kp observations from The World Data Center for Geomagnetism in Kyoto, Japan
- Compiled November 2011-May 2021, 0030 UTC Kp forecasts from NOAA Space Weather Prediction Center (SWPC) through an internal service
- Analyzed data using R statistical software and Verification package
- Used only *whole number* Kp-Index values because finer granularity (i.e. -, o, +) was not readily available for the forecasts (See Table 1).
- Forecasts were treated as categorical or binary.
- Compared 24-hour forecast & observation values in for each 3-hour synoptic period, for daily maximum value and for storm, no-storm forecasts.

References

- Wilks, Daniel S. *Statistical Methods in the Atmospheric Sciences.* Elsevier, San Diego, 2019.
- https://www.7thverificationworkshop.de/Presentation/tutorial_verification_of_forecasts_of_continuous_variables.pdf
- Jolliffe, I. T., and David B. Stephenson. *Forecast Verification: A Practitioner's Guide in Atmospheric Science.* Wiley-Blackwell, Hoboken, NJ;Chichester, West Sussex;, 2012;2011;.
- J. Bartels, The technique of scaling indices K and Q of geomagnetic activity, Ann. Intern. Geophys. Year **4**, 215-226, 1957;
- The RECCS Program is funded by the National Science Foundation (grant number EAR 1757930)

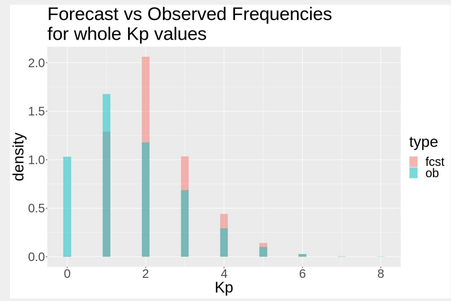

Figure 2: Frequency plot for all Kp values forecast and observed from November 2012 - May 2021. Note the tendency to under-forecast the lowest Kp values (0, 1) and to over-forecast the remaining values.

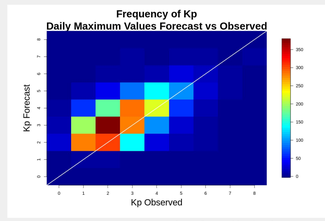

Figure 3: Frequency plot for all *daily maximum* Kp values forecast and observed from November 2012-May 2021. Again, the over-forecasting tendency is evident in this figure. The white diagonal line indicates a perfect forecast.

5. Results & Discussion

Comparison of all 24 hour Kp forecasts and observations (Figure 2) suggests low Kp values are under-forecast, while Kp values above 3 are over-forecast. A similar over-forecasting trend is evident when examining the 24-hour daily maximum Kp forecasts as evidenced by the highest "hit" frequency above the diagonal line. Finally, storm, no-storm forecasts showed a similar tendency with yearly bias scores ranging from +1.27 to +2.1. The monthly Heidke Skill Score (Figure 4) and Threat Score (Figure 5) were also calculated. Improvements that appear to be centered around the equinoxes merit more investigation.

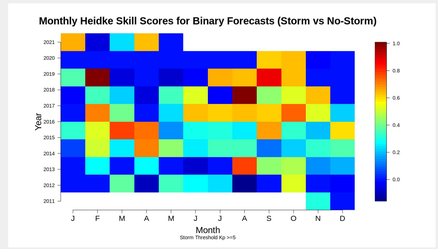

Figure 4: Monthly Heidke Skill Scores for storm, no-storm (i.e. binary) forecasts for November 2012 - May 2021. Here storms are defined as Kp >= 5.

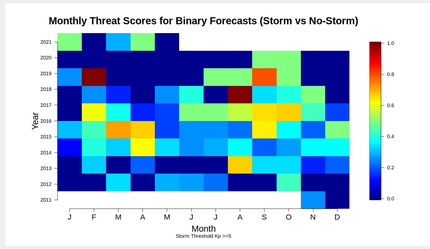

Figure 5: As in Figure 4 but for the Threat Scores.

6. Conclusion.

Initial analysis of verification statistics indicates over-forecasting of the Kp index for Kp=3 (unsettled) and higher indices across all categories. These results would suggest that driving numerical ionospheric or auroral models with Kp forecasts would result in similarly over-forecast phenomena. Significant work remains to thoroughly understand the forecast performance, particularly since the high-impact events are so rare. Interaction with the user community is also required to determine which side of the no-missed-events versus no-false-alarms dichotomy individual users fall.