Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 20 June 2013 (MN LaTeX style file v2.2)

The addition of solar activity (space weather) data to the Open Science Data Cloud in order to facilitate cross-disciplinary studies

Maria T. Patterson[1] , Joshua D. Eisenburg[2], with Rafael Suarez[3]
[1] *New Mexico State University*
[2] *Florida International University*
[3] *University of Chicago, OSDC*

20 June 2013

1 INTRODUCTION

In this paper, we propose the addition of time series solar activity data to the Open Science Data Cloud in order to facilitate data intensive cross-disciplinary discoveries. The light incident on the Earth from the Sun affects the activities of all life on Earth. Our year, defined by the period of rotation of the Earth around the Sun, is the most obvious result of the cyclic effect of the Sun on the Earth. The seasonality of the year is caused by the Earth's axial til and the resulting varying intensity of the *incident* solar flux with latitude on the Earth as it moves along its orbit. However, the solar activity and energy of light *emitted* by the Sun is also changing on various timescales.

While the seasons experienced on Earth have an undeniable effect on the climate and the resulting activities of humans and nature, it is unclear whether there are lower-order effects of variation of solar activity on life on Earth. The Open Science Data Cloud provides an ideal infrastructure for discovering the lower-order effects of variation in solar activity on Earth. With the addition of solar activity data from the Space Weather Prediction Center, we may be able to make discoveries that are important across several disciplines outside of solar physics. We briefly discuss here the following fields in particular: weather and climate research, geoactivity, vegetation studies, public health, and global communication infrastructure.

2 MOTIVATION- VARIABILITY OF SOLAR ACTIVITY

The seasons of weather observed on Earth are due to the variation of incident solar flux at different latitudes on Earth. While this is a direct result of the variation of the location of the *Earth* along its orbit and its axial tilt, the flux of the *Sun* itself also varies over time, both cyclically over a period of approximately 11 years and also in bursts on timescales on the order of minutes. This is an important consideration for lower-order effects on Earth-based activity.

The solar magnetic cycle varies over a period of approximately 11 years, with cycles ranging from about 9 to 14 years (Hale 1908). A new solar cycle begins when the magnetic polarity of the Sun reverses. During the solar cycle, the amount of solar activity (such as sunspots− dark spots caused by intense magnetic activity, flares− brightenings indicating large, protruding magnetic field lines, and coronal mass ejections (CMEs)− massive releases of energetic charged particles) and the average level of irradiation of the Earth from incident solar flux rises and falls.

In Fig. 1 shows the solar magnetic activity cycle over time. Specifically, this graph shows the amount and coverage of sunspots over the surface of the Sun as a function of time. The solar minimum is marked by a sparseness of sunspots and their confinement to only the region of the Sun near the solar equator. The solar maximum is marked by a larger number of sunspots over a larger latitudinal range of the Sun's surface.

Solar activity also varies on much shorter timescales. While individual sunspots may typically survive for many days, solar flares and CMEs persist over a timescale on the order of minutes. These activities affect the level of solar flux and charged particles incident on Earth over short timescales. For example, an increase in aurora borealis and aurora australis (the Northern and Southern lights) seen at high latitudes on Earth is observed following solar flares.

3 RELEVANT DATA SETS

The Space Weather Prediction Center (SWPC) through NOAA maintains a great set of data for interesting for cross-disciplinary research. There are several individual data sets concerning solar activity available that have the potential for high impact on interesting cross-disciplinary research. We describe a few interesting potential data sets briefly below. As of today, all the Space Weather Prediction Center data is now available on the OSDC.

For a long-term analysis of the effects of the solar magnetic cycle, the US National Oceanic and Atmospheric Administration (NOAA) provides a monthly average sunspot count dating from 1749 to the current date. This data would be ideal for comparison to other long-term observations on relatively large timescales (decades).

For analysis on shorter timescales, the Geostationary Operational Environmental Satellites (GOES) observes the solar x-ray and charged particle flux with a time cadence on the order of minutes. We show a plot of the solar x-ray flux for yesterday June 19, 2013 as observed by GOES in Fig. 2.

Additionally, the Coordinated Data Analysis Workshops (CDAW) have compiled a catalog of information for every individual CME event from 1996 to the current date identified from the Large Angle and Spectrometric Coronagraph (LASCO) on the Solar and Heliospheric Observatory (SOHO) satellite. The catalog is available on the CDAW website at http://cdaw.gsfc.nasa.gov/CME_list/.

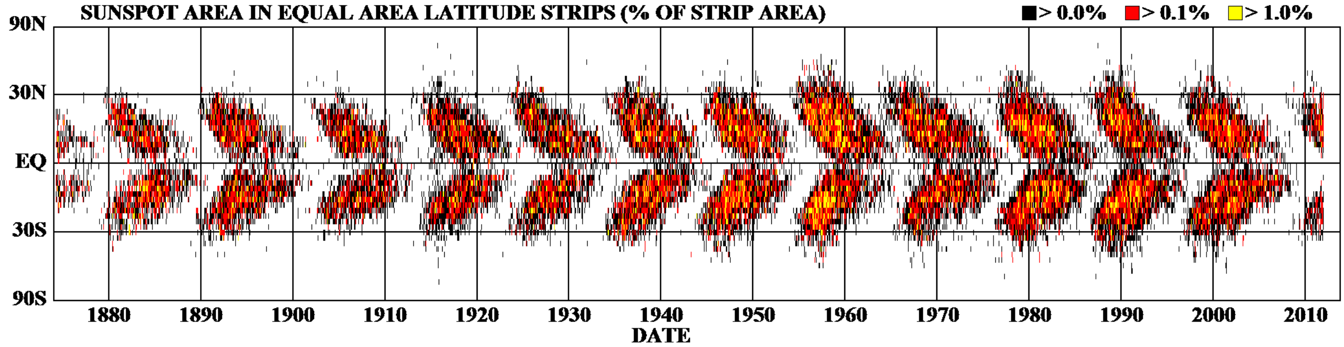

Figure 1. A "butterfly" plot of the sunspot coverage over the surface of the Sun as a function of both latitude on the Sun (vertical axis) and time (horizontal axis) (Hathaway & NASA Marshall Space Flight Center 2013). This graph shows the variation of solar activity over a cycle of roughly 11 years. During minimums in the solar cycle, the few sunspots that are observed are located only near the solar equator. During maximums in the solar cycle, many sunspots are observed across a large latitudinal range of the Sun's surface.

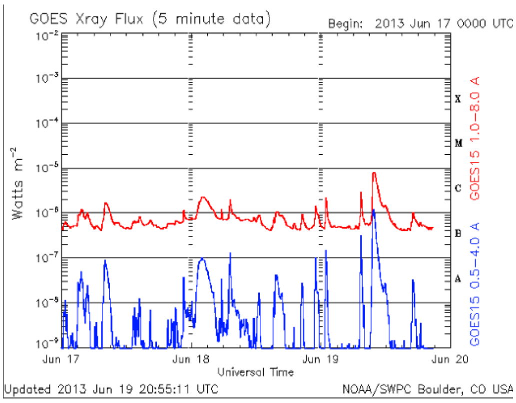

Figure 2. A plot of the solar x-ray flux data from the GOES project as a function of time for June 19, 2013. Red and blue indicate low and high energy x-rays, respectively. (image from the solarmonitor.org)

4 CROSS DISCIPLINARY IMPACT

Solar activity time series data is important for the field of climatology and weather studies, since solar variability likely has a direct effect on Earth's climate. According to the National Research Council's report from January 8, 2013 entitled "The Effects of Solar Variability on Earth's Climate", there is evidence for a correlation between solar maximum and a lower sea surface temperature in the Pacific Ocean. It would be scientifically interesting to look for correlations of the solar activity data with weather and climate information over the entire Earth.

Solar activity data may also be important for studies of vegetation over the Earth. High energy solar flux photodissociates CO_2 in the atmosphere, and we suggest the scientific hypothesis that this may influence vegetation on Earth. Vegetation indices for the Earth as a function of location and time are available from the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument on the NASA Terra (EOS AM) and Aqua (EOS PM) satellites. As a result of discussion of the contents of this paper at this week's workshop, this data has been uploaded to the OSDC.

Seismologists may also be interested in the possible effects of solar activity on seismic activity of the Earth. We suggest the scientific hypothesis that individual solar flares or CMEs could result in an influence of the Earth's geomagnetic field via incident energetic charged particles. This could subsequently result in an variation of seismic activity.

Issues of public health may be interesting to correlate with solar activity. As an interesting research topic, we suggest the hypothesis that variations in the incidence of disease (such as disorders of the skin) over time could be correlated with variations in solar activity over time.

An understanding of space weather is imperative for the management of global communication and satellite systems. A large solar flare caused an hour long radio blackout on Mother's Day of this year. Correlating the strength of solar activity and the incidence of power failures in order to predict and prepare for downtime, for example, would be a very useful project.

5 METHODS AND CHALLENGES

Because solar activity data is in time series, we can compare the pattern of variation in solar activity against any other data of interest that is also in time series by using a time-domain statistical method such as *cross correlation*. A cross-correlation method allows the comparison between two time series data sets offset by a time lag. For example, it is important to consider the time lag between when CMEs are released from the surface of the Sun and when charged particles hit the Earth and the time lag between when photodissociation of CO_2 in the atmosphere occurs and when the effects of vegetation may appear. We would need to take into account the cleanliness of time series data when doing cross-correlation- e.g., how to deal with different sampling rates, smoothing the data in time.

An obvious concern for isolating the effects of solar activity on time series data is removing higher order sources of variation from the data sets. For example, if we are studying the effects of solar activity on the Earth's temperature, we would first want to account for other stronger effects on Earth's climate such as seasonal variability. If we are studying the effects of solar activity on vegetation, we would have to account for other effects such precipitation and overfarming.

6 SUMMARY

We advocate for the exploration of solar activity and space weather data newly added to the Open Science Data Cloud. The solar activity data can be cross-correlated with any relevant time series data set. This data has the potential for high impact discoveries across several disciplines. Here, we have briefly described the impact of solar activity for fields outside of solar astronomy, including climate and weather studies, seismic activity analysis, vegetation studies, issues of public health, and global communication infrastructure.

ACKNOWLEDGMENTS

We acknowledge Zac Flamig for suggesting landcover.org for locating vegetation data that has been uploaded to the OSDC as of today, Dr. Robert Grossman for break time discussions about this topic, and Dr. James McAteer for directing us to the CME event catalog. Rafael Suarez is responsible for the very prompt addition of the MODIS data set and all of the SWPC data to the OSDC.

We acknowledge www.solarmonitor.org for the solar x-ray flux image from yesterday.

REFERENCES

Hale G. E., 1908, ApJ, 28, 315

Hathaway D., NASA Marshall Space Flight Center, 2013, Sunspot coverage area over time: http://solarscience.msfc.nasa.gov/