New Astronomy 62 (2018) 62–69

Contents lists available at ScienceDirect

New Astronomy

journal homepage: www.elsevier.com/locate/newast

Forecasting space weather over short horizons: Revised and updated estimates

Gordon Reikard

Statistics Department, U.S. Cellular, 8410 West Brywn Mawr, Chicago, IL 60631, United States

ARTICLE INFO

Keywords:
Space weather
Forecasting
Frequency domain models
Time series models

ABSTRACT

Space weather reflects multiple causes. There is a clear influence for the sun on the near-earth environment. Solar activity shows evidence of chaotic properties, implying that prediction may be limited beyond short horizons. At the same time, geomagnetic activity also reflects the rotation of the earth's core, and local currents in the ionosphere. The combination of influences means that geomagnetic indexes behave like multifractals, exhibiting nonlinear variability, with intermittent outliers. This study tests a range of models: regressions, neural networks, and a frequency domain algorithm. Forecasting tests are run for sunspots and irradiance from 1820 onward, for the Aa geomagnetic index from 1868 onward, and the Am index from 1959 onward, over horizons of 1–7 days. For irradiance and sunspots, persistence actually does better over short horizons. None of the other models really dominate. For the geomagnetic indexes, the persistence method does badly, while the neural net also shows large errors. The remaining models all achieve about the same level of accuracy. The errors are in the range of 48% at 1 day, and 54% at all later horizons. Additional tests are run over horizons of 1–4 weeks. At 1 week, the best models reduce the error to about 35%. Over horizons of four weeks, the model errors increase. The findings are somewhat pessimistic. Over short horizons, geomagnetic activity exhibits so much random variation that the forecast errors are extremely high. Over slightly longer horizons, there is some improvement from estimating in the frequency domain, but not a great deal. Including solar activity in the models does not yield any improvement in accuracy.

1. Introduction

Forecasts for space weather are used in several areas, ranging from electric power to navigation and communications systems. The horizons are generally short, from as little as a few hours to as long as several days. Geomagnetic series show evidence of being complex processes. There is a clear influence for solar activity. Geomagnetic indexes exhibit cycles at 27 days, corresponding to the rotation of the sun, and the 11–22 year Schwabe and Hale cycles, corresponding to reversals in the sun's magnetic field, although they are slightly dephased (Russell and McPherron, 1978). High magnetic activity during the twentieth century corresponded with a period in which solar activity was at its highest level in 11 millennia (Solanki et al. 2004, Frohlich, 2009). At the same time, there are periods in which the geomagnetic series are not closely correlated with the solar cycle, notably the 1970s (Rangarajan and Iyemori, 1997).

Geomagnetic activity also reflects the rotation of the earth's core, and local currents in the ionosphere (Jordan, 1979; Buffett, 2000; Weiss, 2002). The solid inner core, with a radius of about 1220 km, rotates within a liquid outer core, extending out to 3400 km. The circulation of the outer core is driven by the transfer of heat from the inner core, to the core-mantle boundary. In effect, the magnetic field can be modeled as a dynamo: the change in the magnetic field depends on the diffusivity and the velocity of the fluid. Changes in magnetism generate an electric field, while electric currents give rise to magnetic fields. The magnetic field in the outer core has been estimated as roughly 50 times more intense than at the surface (Buffett, 2010).

The combination of these two processes – solar activity and the internal dynamo – implies that the geomagnetic series may be difficult to predict. The interaction two or more stochastic processes implies that geomagnetic activity should have multifractal properties (Lovejoy and Schertzer, 2013). The solar influence in and of itself may be difficult to forecast beyond very short horizons, since solar activity shows evidence of chaos (Hansen and Willson, 1997; Solanki and Krivova, 2011; Feynman and Ruzmaikin, 2011).

The literature on space weather forecasting includes several methods, ranging from regressions to neural networks and artificial intelligence. The approach tested here includes both frequency and time domain methods, and using solar activity as an input in models for geomagnetic indexes. The findings revise and update earlier results in

E-mail address: Reikarsen@msn.com.

https://doi.org/10.1016/j.newast.2018.01.009
Received 30 November 2017; Received in revised form 8 January 2018; Accepted 17 January 2018
1384-1076/

Reikard (2011, 2013), which were later determined to have been biased by programming errors. The programming errors were not present in and did not affect the results in Reikard (2015).

Section 2 reviews the data. Section 3 sets out the forecasting models. Section 4 runs forecasting experiments. Further analysis is presented in Section 5, and Section 6 concludes.

2. The data

Sunspots have been directly observed since the seventeenth century, and daily sunspot numbers have been reconstructed by interpolating gaps in the actual observations. The sunspots data was obtained from the long-term solar observations website, http://www.sidc.be/silso/datafiles, maintained by the Royal Observatory of Belgium, and supported by the International Council for Science. The dataset used was the revised series developed in Clette et al., (2014), starting in 1818.

Since late 1978, total solar irradiance has been directly measured by satellites, and is available in the ACRIM database (ACRIM, 2014; Willson and Mordvinov, 2003; Scafetta and Willson, 2014). Daily irradiance has been reconstructed to as far back as 1610 (Krivova et al., 2010; Vieira et al., 2011). The values for irradiance were downloaded from a website maintained by a consortium of laboratories and universities in several countries: http://www.mps.mpg.de/projects/sun-climate/data.html. The data set used here is the SATIRE-H series for total solar irradiance, and is updated to 2017 using the ACRIM series. Irradiance is in watts per meter squared (W/m^2).

Figs. 1 and 2 show 30-year cross sections of the sunspot and irradiance data sets. Both of them show several characteristic features, the 27-day cycle, the 11–22 year cycles, and an upward trend in the early part of the twentieth century. Both also show nonlinear variability, with irregular outliers at intermittent intervals.

Two geomagnetic data sets were used. The Aa index, which begins in 1868, is based on the magnetic activity measured at two antipodal stations, Canberra, Australia, and Hartland, England (Mayaud, 1972). The index is the average of the northern and southern values of magnetic activity, weighted to account for differences in the latitudes of the two stations and local induction effects. The Aa activity has the advantage of being the longest record of geomagnetic activity.

The Am and Ap indexes were developed from the K index, a 3-hour measure of activity relative to an estimated quiet day. It was introduced at the beginning of the 1940s (Mayaud, 1980; Menvielle and Berthelier, 1991), and extrapolated backwards using data from selected magnetic observatories. The Ap index, beginning in 1932, uses 13 observatories. The Am index, beginning in 1959, spans 23 locations, 13 northern and 10 southern stations, arranged in groups representing longitudinal sectors (Menvielle and Berthelier, 1991; McPherron, 1995; Mayaud, 1980). Because the Am index spans a wider range of observation points, it is used in the following tests. The data prior to 2011 were downloaded from the National Geophysical Data Center (NGDC, 2014). Data from 2011 onward were obtained from the International Service of Geomagnetic Indexes (http://isgi.unistra.fr/data_download.php). The data run through May 31, 2017.

Figs. 3 and 4 show short cross-sections of the Aa and Am indexes. Both show high degrees of nonlinear variability. During periods of high solar activity, the data is more volatile than at the trough of the solar cycle.

3. The forecasting models

Reviews of the literature are available in McPherron (1995) and Lundstedt (2005). Many of the earlier studies focused on a single class of models. Regression-based methods were used in Baker et al., (1990), Blanchard and McPherron (1993, 1995), Papitashvili et al., (1998), and Codrescu et al., (2004). One limitation of regressions is that they tend to capture the central tendency of the data more effectively than the turning points. In part for this reason, many studies have preferred neural networks. References include Lundstedt (1996), Gleisner and Lundstedt (1997), Wu and Lundstedt (1997a,b), O'Brien and McPherron (2000), Tulunay et al., (2005), Martin et al., (2005), Vandegriff et al., (2005), and Wang et al., (2008).

Other artificial intelligence techniques, including genetic algorithms, have been used in McPherron (1993), Wintoft and Lundstedt (1998), Orfila et al., (2002), Mirmomeni et al., (2006, 2007), Vorotnikov et al., (2008). Mirmomeni et al., (2007) use co-evolving models. Mirmomeni et al., (2010, 2011) and Gholipour et al., (2007) combine neuro-fuzzy techniques with spectral methods (see also Kalhor et al., 2011, 2012). Cao and Cao (2006) combine neural nets and wavelets.

Despite the popularity of more advanced techniques, regressions with time-varying coefficients have often proven very effective in dealing with nonlinear data (Bunn, 2004; Granger, 2008). Let Y_t denote geomagnetic activity, *ln* denote natural logs, ω denote a coefficient, the

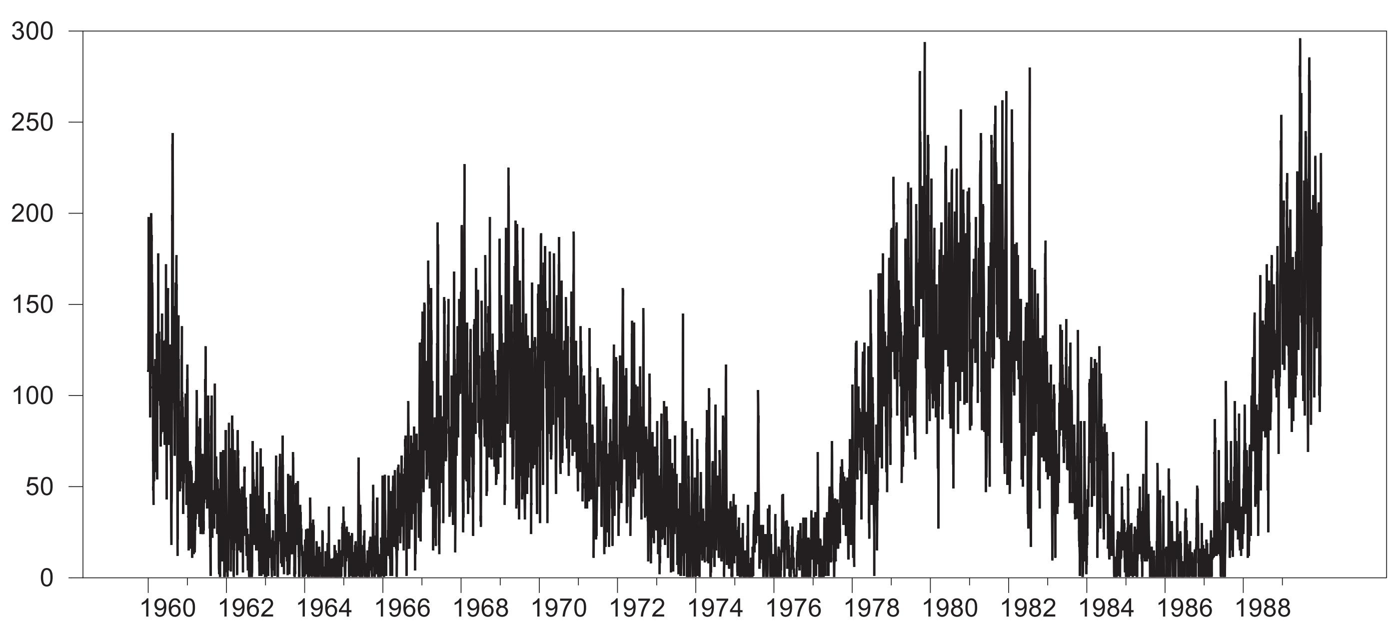

Fig. 1. Daily sunspot number, January 1, 1960 through December 31, 1989.