Physica A 493 (2018) 47–53

Contents lists available at ScienceDirect

Physica A

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

Does solar activity affect human happiness?

Ladislav Kristoufek

Institute of Economic Studies, Faculty of Social Sciences, Charles University, Opletalova 26, 110 00, Prague 1, Czech Republic

ARTICLE INFO

Article history:
Received 14 September 2016
Received in revised form 23 August 2017
Available online 4 November 2017

Keywords:
Happiness
Solar activity
Sunspot numbers
Online data
Twitter

ABSTRACT

We investigate the direct influence of solar activity (represented by sunspot numbers) on human happiness (represented by the Twitter-based Happiness Index). We construct four models controlling for various statistical and dynamic effects of the analyzed series. The final model gives promising results. First, there is a statistically significant negative influence of solar activity on happiness which holds even after controlling for the other factors. Second, the final model, which is still rather simple, explains around 75% of variance of the Happiness Index. Third, our control variables contribute significantly as well: happiness is higher in no sunspots days, happiness is strongly persistent, there are strong intra-week cycles and happiness peaks during holidays. Our results strongly contribute to the topical literature and they provide evidence of unique utility of the online data.

1. Introduction

The influence of solar activity on human behavior and health has not only been a basis of various national sayings and folklore, it has also been an active and controversial scientific topic for decades [1–4]. As reviewed by Lipa et al. [4], the solar activity influences magnetosphere and ionosphere of the Earth both with short-lived disturbances such as solar flares and with long-lived structural changes in the magnetic field. However, other effects of the solar activity on the Earth, its climate, wildlife and human health and behavior have remained controversial until present days.

Mendoza & Diaz-Sandoval [5] study the relationship between solar activity and occurrence of myocardial infarctions in Mexico City and find a weak positive relationship. In their following study, Mendoza & Diaz-Sandoval [6] enlarge the dataset and they find that solar activity maxima are the most hazardous for myocardial infarctions deaths, mainly for age groups above 65 years. Broadening the study outside of Mexico as well, Mendoza & de la Peña [7] focus on lower latitudes, specifically Cuba and low latitudes of Mexico, and confirm the findings even for these conditions.

Cherry [8] argues and empirically shows (on the dataset from southeast Asia) that solar activity is correlated with the Schumann resonances, deviations of which are detected by human brain and lead to health problems and possible excessive death rates through the melatonin mechanism. Palmer et al. [9] review the topical literature focusing on the field of heliobiology. They specifically discuss the Schumann resonances potential to explain the effect of solar activity on human health and, in a light of the reviewed literature, find it a promising link. Babayev & Allahverdiyeva [10] study the effect of geomagnetic variability on brain functioning using EEG and find that abrupt changes in geomagnetic variability have strong negative effect mainly on emotional and vegetative parts of brain while personality characteristics are not affected significantly. Some studies take the effect of solar activity even further and e.g. Mikulecky [11] finds that historical revolutions culminated close to solar activity maxima while flourishing periods were near solar minima.

The recent studies thus suggest that increased solar activity has negative effect on humans, both from physical and mental health perspective. However, most of these studies are fixed in time, either by being connected to a specific event

E-mail address: ladislav.kristoufek@fsv.cuni.cz.

https://doi.org/10.1016/j.physa.2017.10.031
0378-4371/

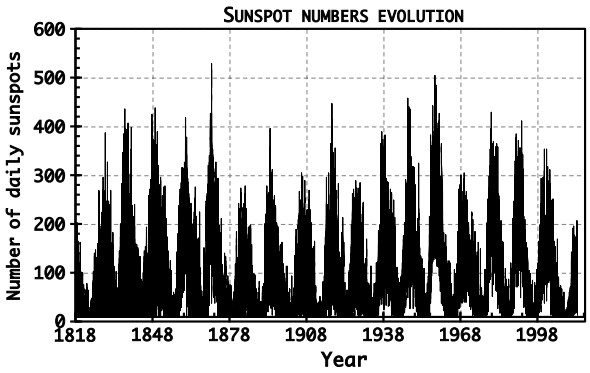

Fig. 1. Evolution of daily sunspot numbers. Daily time series are available down to year 1818. The series follows a strongly cyclical path with a dominant scale of around 11 years.

of an increased solar activity or by being laboratory/experimental. This is mainly due to a problematic availability of data, specifically time series, of evolution of human health, be it physical or mental, that could filter out a possibility that the reported results are due to some other factors which are characteristic for the given event.

Here we focus on the effect of solar activity on human happiness. As the human happiness is generally not easily observable, we make use of the Happiness Index based on sentiment of Tweets (posts on Twitter) for a given day (more details are given in the Data section). Utilizing online data, and specifically the Twitter activity, has already proven useful in various disciplines [12–20]. In our specific case, it allows us to quantify the relation between solar activity, represented by sunspot numbers, and happiness. As both these series are available at daily frequency, it allows us to study the relationship as a standard multivariate time series relationship. In the following section, we provide a detailed description of statistical and dynamic properties of the analyzed series. The next section introduces a set of four models we use. And the last section presents the results with some additional discussion.

2. Data

We work with two base time series—sunspot numbers and the Happiness Index. Sunspot numbers serve as a proxy series for solar activity and it represents daily number of sunspots[1] [21]. The series is obtained from the Solar Influence Data Analysis Centre (SIDC) at http://www.sidc.be/sunspot-data and it is shown in Fig. 1. The daily dataset is available back to 1818. It is evident that sunspot numbers follow a strong cyclical pattern which is well documented in the literature with a dominant cycle of 11 years [22]. There are also days when no sunspot numbers were recorded, which might influence our final regression and we thus need to take this aspect into consideration.

The Happiness Index series is obtained from http://hedonometer.org and it is available from 10.9.2008 to 27.5.2015, which gives 2431 observations (the sunspot numbers series is constrained accordingly). The index is based on randomly sampled Tweets (approximately 10% of all Tweets) for each given day. The happiness for a day is given by language processing of the sampled Tweets with a use of the Amazon's Mechanical Turk on the set of approximately 10,000 most frequently used words based on collections of Google Books, New York Times articles, Music Lyrics, and Tweets. More details about the selection procedure and lists of words can be found in Dodds et al. [23–26]. For each of these words, a value between 1 (sad) and 9 (happy) has been attributed. Based on the Amazon's Mechanical Turk, each of the 10,000 words gets an average happiness value. Then, for each day using the Twitter's Gardenhose feed, 10% of all English Tweets are recorded and aggregated together. The Happiness Index is then a weighted average of all relevant (being one of the most frequently used ones, i.e. with a score) words for a given day with a weight given by the happiness score of the given word.

The procedure has its limitations—the index represents only the Twitter posts in English and there are rare specific cases, which are questionable.[2] Nonetheless, the index is very unique and it provides an irreplaceable source of information (at least at this point[3]).

[1] According to the Solar Influence Data Analysis Centre (SIDC) information, the daily total sunspot number is derived as $N_s + 10 * N_g$ where N_s is the number of sunspots observed and N_g is the number of groups counted over the entire solar disk. More details and history of the measures can be found on the provided webpage.

[2] As a representative, we pick the day of Osama bin Laden's death (2 May 2011). This day is a very sad day based on the index because the Tweets contain words "dead", "death", and "killed" which are considered to be negative. However, we can speculate that this was rather a happy day for the English speaking countries which are represented by English Tweets. Nevertheless, this is one of only few exceptions when the index gives opposite than expected happiness scores.

[3] In an ideal situation, we would be able to obtain the list and content of the Tweets used for the index construction. The index could then be recreated using a more rigorous natural language processing which would have better handled the outlying sentiment values as the ones mentioned in the previous footnote. Unfortunately, these are not available.