

## Correlability of Solar Wind with Seismic Events in the Balkan Peninsula Zone

FILIP ARNAUT<sup>1</sup>, DEJAN VUČKOVIĆ<sup>1</sup>,  
IVANA VASILJEVIĆ<sup>1</sup> & VESNA CVETKOV<sup>1</sup>

**Abstract.** The Solar Heliospheric Observatory (SOHO) satellite was launched on the 2<sup>nd</sup> of December 1995 at L1 Lagrange point ( $1.5 \times 10^6$  km from Earth) with the purpose of gathering data for helioseismology, remote sensing of the solar atmosphere, and solar wind *in situ*. The satellite was positioned into orbit in early 1996, with data acquisition expected to commence on January 20<sup>th</sup>. The correlation between increased values of solar wind parameters and earthquakes in the Balkan peninsula zone between 1996 and 2018 was made possible by data obtained through continuous proton density and proton velocity monitoring. The assessment of the anomalous threshold was based on statistically determined parameters due to the huge fluctuation of solar wind over time and distinct value increases of proton density and speed. Visual representations of proton density and proton speed were created for the time window preceding each earthquake after defining the boundary between normal and anomalous values. According to the chart analysis, increased proton density occurred in 40 of the 50 cases observed, whereas increased proton velocity appeared in 28 of the 50 cases. Using hypergeometrical probability and an unbiased test with randomly generated parameters, the discovered correlation was statistically verified. A retrospective selection bias analysis is also provided in the research paper.

### Key words:

*earthquake prediction,  
solar wind, statistical validation  
of earthquake precursors.*

**Апстракт.** Сателит соларне хелиосферске опсерваторије (SOHO) лансиран је ка L1 Лагранжовој тачки ( $1.5 \times 10^6$  km удаљена од Земље) 2. децембра 1995. године са циљем прикупљања података за потребе хелиосеизмологије, даљинске детекције соларне атмосфере и соларног ветра *in situ*. Сателит је доспео у орбиту почетком 1996. године са почетком аквизиције података 20. јануара. Подаци прикупљени континуалним мониторингом густине и брзине протона омогућили су корелацију појава повећаних вредности параметара соларног ветра и потреса на Балканском полуострву у периоду од 1996. до 2018. године. Због велике варијабилности соларног ветра у времену и изразитих скокова густине и брзине протона, одређивање аномалије густине и брзине протона рађено је на основу статистички срачунатих параметара. Након одређивања границе нормална вредност/аномалија, конструисани су графици густине и брзине протона за референтни период пред

<sup>1</sup> University of Belgrade, Faculty of Mining and Geology, Department of Geophysics, Đušina 7, 11000 Belgrade, Serbia.  
E-mail: G601-20@rgf.bg.ac.rs

**Кључне речи:**

*предвиђање потреса,  
соларни ветар, статистичка  
валидација узрочника потреса.*

сваки потрес. Анализа графика показала је да се појава повећане густине протона десила у 40 од 50 случајева, а повећана вредност брзине протона догодила се 28 од 50 случајева. Добијена корелација је статистички верификована хипергеометријском вероватноћом и независним тестом са насумично генерисаним параметрима. У раду је приказана и анализа ретроспективне пристрасности избора која се манифестује приликом истраживања прекурсора потреса.

**Introduction**

The term “precursor” refers to a wide variety of physical phenomena that are used for earthquake prediction (CICERONE et al., 2009). Earthquake prediction is a branch of seismology that seeks to predict forthcoming earthquakes in the short, medium, and long term. Deterministic approach in earthquake prediction requires the prior knowledge of the earthquake epicenter’s geographical latitude and longitude, magnitude, and time, which can be seen as unrealistic compared to a probabilistic approach (similar to weather forecasting). The probabilistic approach yields a likelihood that an earthquake will occur in a given region at some time span. Every earthquake prediction method should be based solely on statistics i.e., probability (KAMER et al., 2021).

In order for earthquakes to be predicted with high certainty, the search for new precursors is ongoing still to this day. Today, one group of precursors are astronomical precursors. Modern research showed that there is a strong correlation between solar wind parameters (density, velocity, dynamic proton pressure, etc.) and global earthquakes with a minimum magnitude of M5.6 (MARCHITELLI et al., 2020). A positive correlation was also displayed with earthquakes that occur on a global scale and an increase in proton density (STRASER & CATALDI, 2014). Aside from proton density, vertical Z Earth’s magnetic field component was positively correlated with earthquakes with a magnitude M6.0 (STRASER & CATALDI, 2014).

It is also worth noting that increased solar wind parameter values in the days leading up to the earthquake can’t be considered as a precursor. Earthquake precursors are physical phenomena that occur as a result of an unstable subsurface state

in a particular region. As a result, precursors are the result of accumulated stress. Increases in solar wind parameters can only be considered possible triggers. Increased solar wind parameter values can act as a “straw that broke the camel’s back” by accelerating up an earthquake that would have happened anyhow and was tectonically controlled because of the tectonically accumulated stress (MULARGIA, 1997; MULARGIA, 2001)

Even if the increased solar wind parameters do have an effect on tectonically controlled earthquakes acceleration, it is crucial to understand the modulation they cause. Proposed mechanisms of the Sun’s influence on earthquakes should be discussed to get a better knowledge of those modulations.

SIMPSON (1967) presented the first mechanism, which states that the magnetohydrodynamical interaction of the solar and terrestrial magnetic fields can affect the Earth’s angular velocity. Within that mechanism SIMPSON (1967) states there are two sub-mechanisms that can alter the Earth’s angular velocity, causing earthquakes. The first sub-mechanism states that the continents and oceans are dynamically unstable. Variations in the rotating velocity of the Earth can produce changes in the subsurface stress level (SIMPSON, 1967). The second sub-mechanism states that the Earth’s viscosity prevents it from adapting to the angular velocity change. Inability to adapt to a new state might produce tension in the upper part of the Earth’s crust, causing an earthquake.

SIMPSON’S (1967) second mechanism is similar to MARCHITELLI et al. (2020) reported mechanism in certain ways. Electrical (telluric) currents passing through the subsurface can induce stress pulls due to the reverse piezoelectric effect, which, along with accumulated tectonic stress, can further destabilize

the fault and cause an earthquake. SIMPSON (1967) claimed that telluric currents raise subsurface temperatures, which destabilize faults. However, the subsequent study revealed that this claim was erroneous, as the subsurface temperature increase was two orders of magnitude lower than it should be to induce an unstable subsurface condition. MARCHITELLI et al. (2020) claimed that their mechanism was tectonically controlled, but that an external trigger was used (telluric currents increased by the solar wind).

During the construction of a precursory hypothesis, the main goal is to make sufficient advancement compared to the already existing precursory hypothesis. Because of that, a balance should be made between generalities and particularities (RHOADES & EVISON, 1989; MULARGIA, 1997). Prior mentioned research (particularly MARCHITELLI et al., 2020 and STRASER & CATALDI, 2016) gave a foundation for a statistical analysis of the correlation between solar wind parameters and earthquakes that occur in a region of complex (seismo)tectonic architecture, such as the Balkan peninsula.

## Methodology

The CELIAS Proton Monitor (solar wind proton density and velocity data) and the United States Geological Survey (USGS) Earthquake catalog were used in this study. These two databases were used for a period of time ranging from 1996 to 2018.

The proton density and velocity datasets contain values measured every 30 seconds, i.e., a dataset of 1051200 datapoints is regarded full for a year with 365 days, while a dataset of 1054080 datapoints is considered full for a leap year (366) days. The true datapoint number in each dataset was computed, and the results ranged in the order of 90 percent capacity, except for 1998, where the datapoint capacity in the dataset is 59 percent. The explanation for the discrepancy is that the satellite lost control, lost power, and was no longer directed at the Sun from June 24 until October 29. The satellite was put back into operation towards the end of October 1998.

From 1996 to 2018, the USGS website provided an earthquake catalog at a worldwide scale with a minimum magnitude of M5.0. Datasets containing

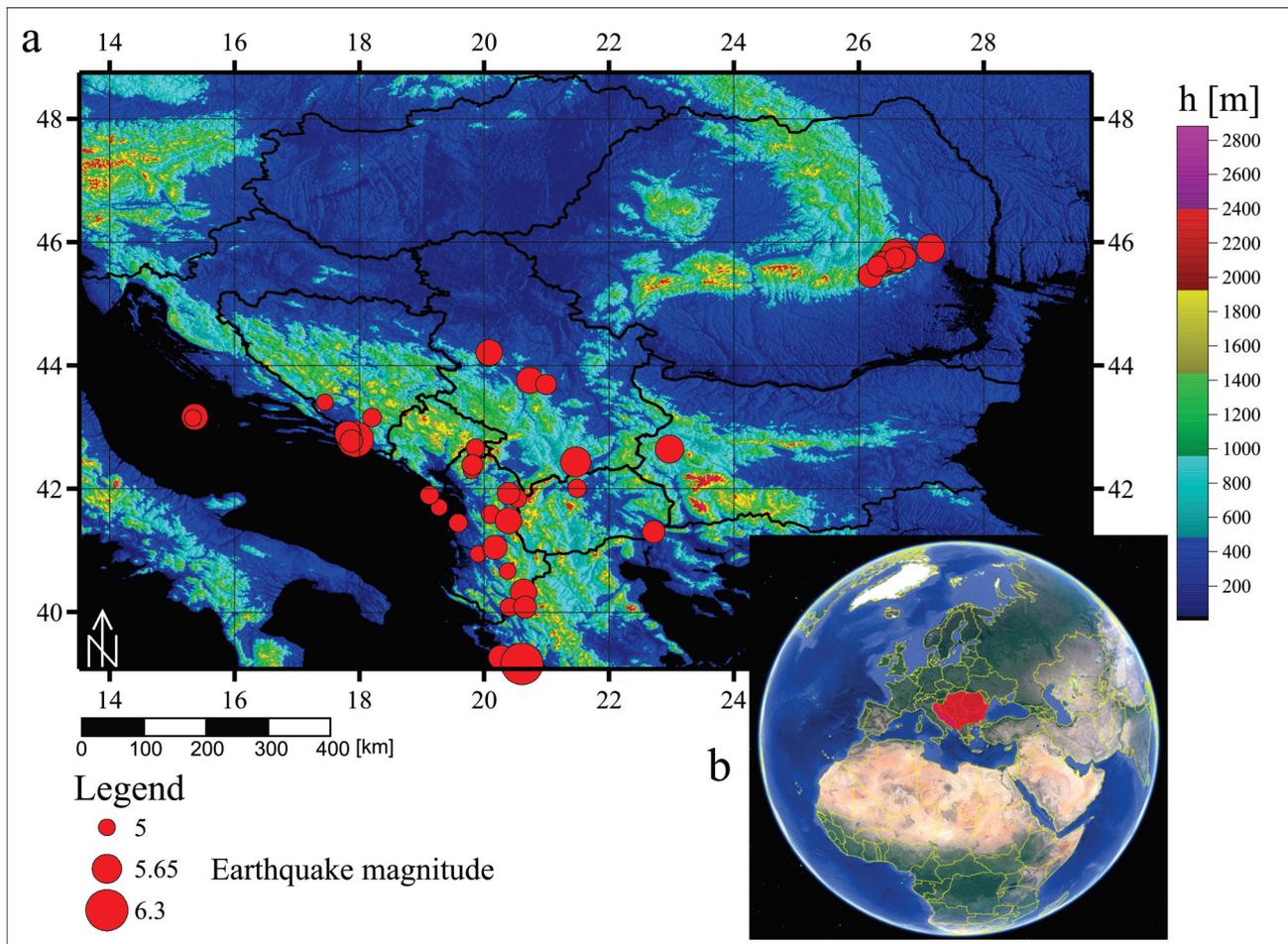
information about earthquakes in the Balkan peninsula zone (mainly Dinarides, Carpatho-Balkanides, and the Pannonian Basin) were selected from the worldwide datasets. In the Balkan Peninsula region, 52 earthquakes with a minimum magnitude of M5.0 occurred over a 23 year period. Only 50 earthquakes were considered in the final analysis since two earthquakes that occurred on the territory of the Republic of Serbia occurred during a period when the satellite did not acquire any data. Three earthquakes that are thought to have occurred on the border between Albania and Greece were also taken into account. It should also be mentioned that the Republic of Hungary's territory was assumed to span a larger region of the Pannonian basin, despite the fact that no earthquakes with a minimum magnitude of M5.0 occurred on the territory from 1996 to 2018. Aside from the earthquakes already described, three more earthquakes in the Adriatic Sea region have been added.

It is worth noting that the earthquakes were simply filtered from the USGS database by the name of the country in which they occurred. The main goal was to cover as much of the Balkan peninsula as possible, including the Dinarides, Carpatho-Balkanides, and Pannonian Basin, which are three separate geological units in the Balkan Peninsula.

The digital elevation model (DEM: see references NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION, for the internet link) base can be used to map the earthquakes outlined previously (Fig. 1). The DEM utilized was GLOBE, which has a 1 km resolution. Figure 1a shows that the majority of earthquakes occur in Albania (Dinarides) and Romania (Carpatho-Balkanides), accounting for 68 percent of all earthquakes (34/50). The area of the researched surface in relation to the Earth's surface can be observed in figure 1b. The examined area is only 0.41 percent of the overall Earth's size, with a total area of 2 092 846 km<sup>2</sup>.

The analysis was carried out with a total of six steps which will be discussed below:

**Step 1:** Basic statistical calculation for obtaining the information about the mean value, standard deviation, minimum, and maximum values for the yearly dataset. This type of data provides a foundational understanding of the dataset.



**Fig. 1. a.** Location of earthquakes with a DEM base; **b.** Size comparison of the investigated area and Earth's area (modified after: <https://earth.google.com/web/>)

**Step 2:** Determination of the anomaly threshold, which is the line between what is considered normal and what is considered anomalous. This is done by calculating skewness and kurtosis (and two standard errors of skewness and kurtosis), as well as basing the anomalous threshold on  $n$  standard deviations. According to the “33/67/99.7” rule (empirical rule used to estimate the percentage of data that falls within one, two, or three standard deviations), three standard deviations for normal distributions can be considered anomalous since they correspond with only 0.3 percent of data. The  $n$  is greater for distributions with larger skewness. This stage calculates the percentage of data that is equal to or greater than the anomaly threshold; this information is used to check the quality of the determined anomaly threshold. That value should not exceed 1% on average throughout the whole 23 year sample.

**Step 3:** Construction of graphic representation (charts) for the two weeks leading up to each earthquake.

**Step 4:** Identifying which earthquakes have a proton density (velocity) anomaly in the two-week, one-week, and four-day periods prior to the earthquake.

**Step 5:** Hypergeometrical probability based statistical significance test. In this stage, the number of unique anomalous days for each year is calculated, and the hypergeometrical probability is calculated to see if the anomaly occurred more frequently than might randomly. In addition, an independent test is run to corroborate the hypergeometrical probability as a statistical significance test.

**Step 6:** Analysis and interpretation of the results from the previous five steps. The correlation is thought not to be significant, i.e., can be thought to

be random, if the probability obtained by the hypergeometrical probability coincides in the range of 10% with the number of times analyzed earthquakes had an anomalous proton density (velocity) value in the time window chosen.

## Results and Discussion

In a yearly format, Table 1 shows general statistical information of the proton density parameter. The mean value in the database (from 1996 to 2018) shows a decreasing trend until 2008, after which there is a modest increase until 2015, after which the mean yearly proton density decreases again (Fig. 2, middle). The average number of protons per cubic centimeter over the entire dataset is 5.3.

**Table 1.** General statistical information for the proton density dataset.

Year	Mean [ $\text{cm}^{-3}$ ]	Standard deviation [ $\text{cm}^{-3}$ ]	Minimum [ $\text{cm}^{-3}$ ]	Maximum [ $\text{cm}^{-3}$ ]
1996	7.82	4.43	-1	53.01
1997	8.65	4.96	-1	57.65
1998	7.67	4.61	-1	66.22
1999	6.35	4.52	-1	57.28
2000	6.31	4.82	-1	70.18
2001	6.50	5.38	-1	120.77
2002	6.96	5.30	-1	70.56
2003	5.38	3.98	-1	72.19
2004	4.97	3.87	-1	111.81
2005	4.95	4.58	-1	88.56
2006	4.68	4.24	-1	75.65
2007	4.06	3.99	-1	82.66
2008	3.74	3.08	-1	72.44
2009	4.03	3.14	-1	71.32
2010	4.35	3.53	-1	117.17
2011	4.33	3.87	0	90.81
2012	4.57	4.02	0	118.02
2013	4.45	4.00	0.01	126.92
2014	4.81	3.49	0.06	60.29
2015	5.27	4.43	0.07	77.81
2016	4.55	3.54	0.09	88.19
2017	3.94	3.14	0.06	63.2
2018	3.66	2.86	0.04	62.06
Average	5.30	4.08	/	81.51

With the exception of 2018, which has a standard deviation of  $2.86 \text{ cm}^{-3}$ , standard deviation values are consistent, ranging from 3 to 5 protons per cubic centimeter. The overall dataset's mean standard deviation is  $4.08 \text{ cm}^{-3}$ .

For the time span 1996 to 2010, the minimum values are  $-1$ . This is not a representation of the measured proton density, but rather a lack of data (mainly of technical nature). Minimum values after 2010 are either 0 or a number greater than 0. Values less than 0 were filtered for the period up to 2010.

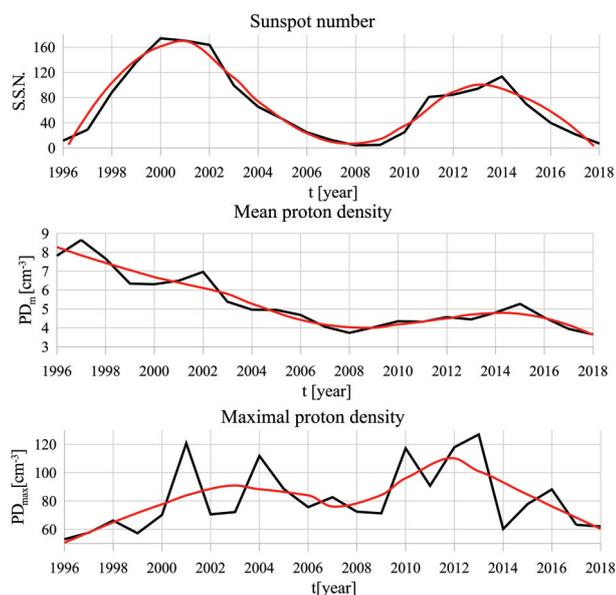
The maximum values in each dataset range from 53 (1996) to 127 (2013) protons per cubic centimeter. With two maximums (2003 and 2012) and three minimums (1996, 2007, and 2018), there is some type of pattern in the maximum levels (Fig. 2, bottom).

Because of the displayed periodicity of the mean and maximal proton density values, a correlation with the number of Sunspots in the same period was performed (Fig. 2, top). To visually represent the Sunspot number, a database of Sunspot numbers was taken (see references SILSO data, Royal Observatory of Belgium, Brussels for the internet link) and a local regression (LOESS) filter was applied.

For the period 1996 to 2018, the sunspot chart (Fig. 2, top) shows three minimums and two maximums, indicating two cycles (Solar cycles 23 and 24). Minimum values can be recorded in 1996, 2008, and 2018 (on average every 11 years), whereas maximum values are also located at an 11 year time period.

Two correlation coefficients, Pearson's and Spearman's, were calculated to assess the correlation between these three factors. Pearson's correlation coefficient is a traditional correlation coefficient that can occasionally reveal flaws, such as when one data point appears to be outside the general trend

of other data points. Pearson's correlation coefficient will yield a lower value than the true correlation coefficient (SCHÖBER et al., 2018). To acquire more detailed information, both correlation coefficients were calculated, and a description and a color based on the



**Legend**  
 — - Measured value  
 — - Filtered value  
 S.S.N.- Sunspot number  
 PD<sub>m</sub> - Yearly mean proton density  
 PD<sub>max</sub> - Yearly maximal proton density

**Fig. 2.** Sunspot number (top) correlation with yearly mean proton density values (middle); Yearly maximal proton density values (bottom)

value range were assigned, as reported by SCHÖBER et al. (2018).

Table 2 shows that maximum proton density values have a weak correlation with the Sunspot number calculated using both Pearson’s and Spearman’s correlation coefficients. On the other hand, the Pearson correlation coefficient shows a weak correlation with annual mean proton density values, whereas

**Table 2.** Correlation coefficients for the solar wind parameters and Sunspot number.

/	Pearson	Spearman	Description	Color	Range
		Sunspot number		Negligible	Red
PD <sub>m</sub>	0.39	0.57	Weak	Blue	0.1-0.39
			Moderate	Yellow	0.4-0.69
			Strong	Green	0.7-0.89
PD <sub>max</sub>	0.16	0.12	Very strong	Dark Green	0.9-1.0

the Spearman correlation coefficient shows a moderate value.

To determine the anomalous threshold, the skewness, kurtosis, and two standard errors of skewness and kurtosis must be calculated first (table 3). From 1.76 in 1996 to 5.34 in 2007, the skewness parameter shows a wide range of fluctuation. Skewness takes an average 3.37 value over the course of the observed period. Because the skewness parameter informs us about the “length of the tail” of the distribution, the skewness is minimal for the lowest maximum value of yearly proton density. The contrary is not true, skewness is not maximal for the highest maximum values in the dataset (2013), but it does take values of around 4. The kurtosis parameter has a wide range of values, ranging from 5.32 in 1996 to 51.58 in 2007. Kurtosis has an average value of 24.31.

**Table 3.** Skewness, kurtosis and two standard errors of skewness and kurtosis for the investigation period.

Year	Skewness	Kurtosis	Year	Skewness	Kurtosis
1996	1.76	5.32	2008	3.73	25.63
1997	2.18	8.56	2009	3.36	23.79
1998	1.89	6.49	2010	4.36	50.29
1999	2.15	7.71	2011	4.30	34.65
2000	2.69	11.87	2012	4.09	40.47
2001	2.98	15.03	2013	3.99	44.35
2002	2.80	13.26	2014	2.52	12.12
2003	3.76	27.41	2015	3.53	20.75
2004	3.38	31.06	2016	3.36	19.82
2005	3.64	23.09	2017	4.14	32.74
2006	3.61	21.79	2018	3.89	31.25
2007	5.34	51.58	<b>Average</b>	3.37	24.31

Reference Skewness Kurtosis  
 1.02 2.04

Average values of skewness and kurtosis are higher than the approximated two standard error values of skewness and kurtosis for the investigated period. This information leads us to the conclusion that the proton density distribution is highly horizontally and vertically distorted, as was verified graphically in figure 3. Because of that, higher multiplicities of standard deviations can be used as the anomaly threshold for each year. For this research, five standard deviations will be used as the anomaly threshold. The anomaly threshold for each year will thus be calculated as:

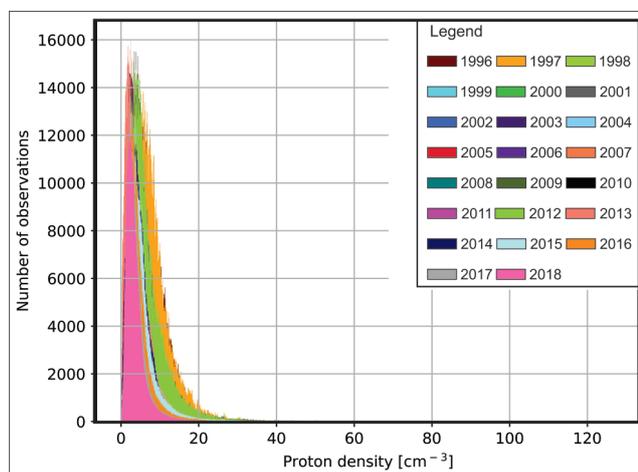
$$Pd_{anomaly}(y) = PD_m(y) + 5SD(y) \quad (1)$$

where:

$PD_{\text{anomaly}}(y)$  - anomaly threshold for each year,

$PD_m(y)$  - mean proton density value for each year,

$SD(y)$  - standard deviation calculated for each year.



**Fig. 3.** Proton density histogram for the period from 1996 to 2018

It is possible to calculate the anomaly threshold and the percentage of data that is equal to or greater than the anomaly threshold using the previously displayed data. The maximal anomaly threshold can be obtained for the year with a comparatively high yearly mean proton density and high standard deviation, since the anomaly threshold is solely determined by the mean yearly proton density value and the standard deviation value. The anomaly threshold values range from 17.97 protons per cubic centimeter (2018) to 33.46 protons per cubic centimeter (2002), as shown in table 4. The overall dataset's mean anomaly threshold value is 25.69 protons per cubic centimeter. The percentage of data that is equal to or higher than the anomaly threshold varies between 0.24 percent (1996) to 0.67 percent (2015). The average percentage of data over the anomaly threshold is 0.49 percent, which is lower than the expected value of 1%.

**Table 4.** Anomaly threshold information for the proton density parameter.

Year	Mean [cm <sup>-3</sup> ]	Standard deviation [cm <sup>3</sup> ]	5 $\sigma$ [cm <sup>-3</sup> ]	Maximum [cm <sup>-3</sup> ]	Percentage of data $\geq 5\sigma$ [%]
1996	7.82	4.43	29.98	53.01	0.24
1997	8.65	4.96	33.43	57.65	0.37
1998	7.67	4.61	30.72	66.22	0.26
1999	6.35	4.52	28.93	57.28	0.41
2000	6.31	4.82	30.42	70.18	0.45
2001	6.50	5.38	33.40	120.77	0.56
2002	6.96	5.30	33.46	70.56	0.52
2003	5.38	3.98	25.30	72.19	0.47
2004	4.97	3.87	24.31	111.81	0.42
2005	4.95	4.58	27.87	88.56	0.63
2006	4.68	4.24	25.87	75.65	0.58
2007	4.06	3.99	24.03	82.66	0.62
2008	3.74	3.08	19.14	72.44	0.61
2009	4.03	3.14	19.74	71.32	0.45
2010	4.35	3.53	21.98	117.17	0.48
2011	4.33	3.87	23.66	90.81	0.65
2012	4.57	4.02	24.65	118.02	0.52
2013	4.45	4.00	24.45	126.92	0.40
2014	4.81	3.49	22.26	60.29	0.47
2015	5.27	4.43	27.43	77.81	0.67
2016	4.55	3.54	22.26	88.19	0.58
2017	3.94	3.14	19.66	63.2	0.54
2018	3.66	2.86	17.97	62.06	0.48
<b>Average</b>	<b>5.30</b>	<b>4.08</b>	<b>25.69</b>	<b>81.51</b>	<b>0.49</b>

**Table 5.** Anomalous proton density occurrences in the time period before every earthquake.

Earthquake location	Year	Month	Day	Anomaly[14 days]	Anomaly [7 days]	Anomaly [4 days]
Serbia	2010	November	3	•	•	•
	2002	April	24	•	•	
	1999	July	1	•	•	
	1999	April	30	•		
Bosnia and Herzegovina	2005	September	27	•	•	•
	2004	May	23	•	•	•
	1996	September	5			
Croatia	1996	September	17	•	•	
	1996	September	9			
Montenegro	2018	January	4	•	•	•
	2014	December	29	•	•	•
Albania	2018	August	11	•	•	•
	2018	July	4	•	•	•
	2014	May	19	•		
	2009	September	6	•	•	•
	2009	August	21	•	•	•
	2007	June	29	•	•	•
	2006	October	19	•		
	2006	June	21			
	2005	July	10	•	•	•
	2004	November	23	•	•	•
	2004	April	7	•	•	•
	2003	August	14			
	2001	April	9	•		
	1999	December	22			
	1997	May	16	•	•	•
1996	August	20	•			
1996	August	5	•			
1996	July	26				
North Macedonia	2016	September	11			
	2009	May	24	•	•	•
Bulgaria	2012	May	22	•	•	•
Romania	2018	October	28	•	•	
	2016	December	27	•		
	2016	September	23	•	•	
	2014	November	22	•		
	2013	October	6	•	•	•
	2009	April	25	•	•	•
	2005	June	18	•	•	
	2005	May	14			
	2004	October	27	•	•	•
	2002	November	30	•		
	2001	July	20			
	2001	May	24	•	•	•
	2000	April	6			
	1999	April	28	•		
1998	March	13	•	•	•	
Adriatic sea	2004	November	25	•	•	
	2003	March	29	•	•	•
	2003	March	27	•	•	•
<b>Sum •/50 earthquakes</b>				40	31	23
<b>Sum • [%]</b>				80	62	46

After calculating the anomaly threshold, determining whether earthquakes had anomalous proton density values in the two weeks, one-week, and four-day periods leading up to it was done. Table 5 shows the individual earthquakes examined in this study, as well as the occurrence of the anomalous proton density value in the time interval before it denoted by the symbol „●“.

Table 5 shows that in the two-week period leading up to an earthquake, 80 percent of earthquakes (40/50) exhibit anomalous proton density values. This number lowers to 62 percent, or 31 out of 50 earthquakes, in the week leading up to an earthquake. For the four days leading up to each analyzed earthquake, 23/50 (46 percent) showed a proton density anomaly.

The first step in determining the statistical significance of these findings is to establish the number of individual anomalous days in a calendar year. To begin, a definition of an anomalous day should be established:

30 seconds and another with a duration of several hours are not the same. In keeping with the survey's generalities, it is considered that every anomaly has the same relevance, regardless of its duration.

After determining the number of individual anomalous days for each year in the dataset, the hypergeometric probability can be calculated, which is the probability that in a year with 365 days and, for example, 22 individual anomalous days, at least one anomalous day can be found in the chosen time window (in this case 14, 7 and 4 days prior to each earthquake). This probability shows the likelihood of finding a proton density anomaly in the time window chosen before each earthquake at random.

To validate the hypergeometric probability accuracy, or whether it is a good probability model for probability computation, an independent test was created with the same input parameters as the hypergeometric probability, with the exception that all of the parameters are picked at random. Figure 4 depicts the flowchart for such a test.

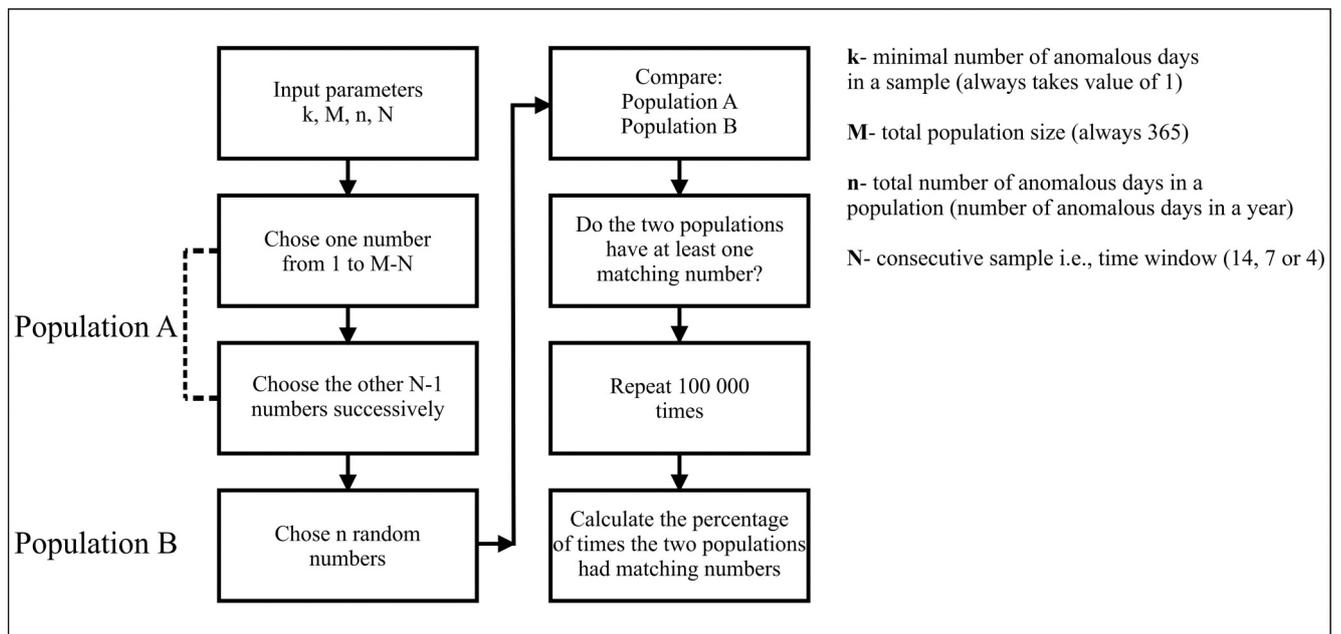


Fig 4. Flowchart of the independent test

*“Every day with at least one proton density measurement higher than the anomaly threshold value for a particular year is an anomalous day.”*

This notion of an anomalous day has some disadvantages. The first is that it ignores the duration of the anomaly, i.e., one anomaly with a duration of

There are eight steps in the flow of the test probability. The first step is to choose a number at random from 1 to M-N (i.e., 365-14/7/4). The remaining N-1 numbers are calculated by adding to the first number +1, +2, +3, etc. This group of numbers is referred to as “Population A,” and it repre-

sents the time period preceding the earthquake. The number of anomalous days in a year is represented by Population B, which is picked at random from 1 to 365 integers. After both Population A and Population B have been determined, the two populations can be compared to see if they have the same numbers. After iterating the method 100 000 times, the percentage of times the two populations had at least one matching number was calculated.

This percentage is an approximation of the hypergeometric probability and is used to check the quality of the probabilities obtained. Table 6 shows the hypergeometric probability, the test probability based on the previously shown algorithm, and the difference between these two probabilities for the time windows of 14, 7, and 4 days before the earthquake.

Table 6 shows that the average hypergeometric probability values for the time windows of 14, 7 and 4 days are 82.52 percent, 59.11 percent, and 40.32 percent, respectively. These probabilities are in line with the algorithm’s probability, which is in the 2 percent range (maximal discrepancy being 2.26 percent for the time window of 7 days). Even if the difference between the two probabilities is small, it can be explained in two ways.

The first reason for the disparity is that as the number of algorithm iterations approaches infinity, the algorithm probability tends to hypergeometric probability. By increasing the number of iterations, this effect can be reduced.

The hypergeometric probability is a cumulative probability of having all anomalous days in a time window sample, which is the second cause of the tiny disparity between the hypergeometric probability and the algorithm probability. Even though the chance of having more than two anomalous days in the time sample is small, it is accounted for in the hypergeometric probability.

With the two explanations in mind, it is easy to see why there’s such a modest difference between the hypergeometric and algorithm probabilities. As a result, the hypergeometric probability model is considered to be a satisfactory probability model for this study.

The obtained hypergeometric probability was compared to the calculated percentage of proton density anomaly occurrence for the time window chosen before the analyzed earthquakes on the Balkan peninsula zone (table 7). If the two values are within 10% of each other, any occurrence of proton density anomaly in the time window preceding an earthquake is regarded to be random. As a result, there is no statistically significant correlation between the two events.

The same method was applied for the proton velocity parameter, with anomalous proton velocity values found in 56 percent (28/50) of earthquakes in the two-week time frame prior to it, 32 percent (16/50) in the one-week time window, and 18 percent (9/50) in the four-day time window. Similar to

**Table 6.** Hypergeometric probability, algorithm probability and difference for the proton density parameter.

General		14- Days			7- Days			4- Days		
Year	Number of anomalous days	H. Probability [%]	A. Probability [%]	Difference [%]	H. Probability [%]	A. Probability [%]	Difference [%]	H. Probability [%]	A. Probability [%]	Difference [%]
1996	22	58.80	57.79	1.01	35.53	34.71	0.82	22.10	21.54	0.56
1997	49	87.23	85.23	1.99	63.87	61.34	2.53	43.96	41.62	2.34
1998	21	57.05	56.00	1.05	34.19	33.61	0.58	21.18	20.69	0.49
1999	29	69.31	67.731	1.57	44.26	42.96	1.30	28.29	27.46	0.84
2000	37	78.24	76.70	1.54	52.99	51.01	1.98	34.91	33.81	1.10
2001	42	82.53	80.80	1.73	57.82	55.65	2.18	38.81	37.17	1.64
2002	38	79.17	77.31	1.86	53.99	52.27	1.73	35.70	34.46	1.24
2003	31	71.81	70.74	1.07	46.56	45.25	1.31	29.99	29.04	0.95
2004	44	84.01	81.99	2.02	59.63	57.10	2.54	40.32	38.56	1.75
2005	52	88.86	86.78	2.08	66.23	63.63	2.60	46.07	43.74	2.33
2006	42	82.53	80.80	1.73	57.82	55.65	2.18	38.81	37.17	1.64
2007	46	85.38	83.41	1.97	61.38	58.82	2.56	41.80	39.87	1.93
2008	51	88.34	86.50	1.84	65.46	62.85	2.61	45.38	43.31	2.06
2009	56	90.73	88.97	1.76	69.16	66.23	2.94	48.79	46.11	2.68
2010	44	84.01	81.99	2.02	59.63	57.10	2.54	40.32	38.56	1.75
2011	46	85.38	83.41	1.97	61.38	58.82	2.56	41.80	39.87	1.93
2012	60	92.31	90.38	1.93	71.88	68.81	3.07	51.40	48.27	3.13
2013	52	88.86	86.78	2.08	66.23	63.63	2.60	46.07	43.74	2.33
2014	47	86.02	84.33	1.69	62.23	60.09	2.14	42.53	40.51	2.01
2015	46	85.38	83.41	1.97	61.38	58.82	2.56	41.80	39.87	1.93
2016	52	88.86	86.78	2.08	66.23	63.63	2.60	46.07	43.74	2.33
2017	52	88.86	86.78	2.08	66.23	63.63	2.60	46.07	43.74	2.33
2018	66	94.21	92.59	1.62	75.56	72.21	3.36	55.13	51.99	3.15
<b>Average</b>	44.57	82.52	80.75	1.77	59.11	56.86	2.26	40.32	38.47	1.85

**Table 7.** Comparison of the hypergeometric probability and the results obtained by the number of proton density occurrences before earthquakes in the Balkan peninsula zone.

Year	Anomaly probability [14 days] [%]	Anomaly probability [7 days] [%]	Anomaly probability [4 days] [%]
1996	58.80	35.53	22.10
1997	87.23	63.87	43.96
1998	57.05	34.19	21.18
1999	69.31	44.26	28.29
2000	78.24	52.99	34.91
2001	82.53	57.82	38.81
2002	79.17	53.99	35.70
2003	71.81	46.56	29.99
2004	84.01	59.63	40.32
2005	88.86	66.23	46.07
2006	82.53	57.82	38.81
2007	85.38	61.38	41.80
2008	88.34	65.46	45.38
2009	90.73	69.16	48.79
2010	84.01	59.63	40.32
2011	85.38	61.38	41.80
2012	92.31	71.88	51.40
2013	88.86	66.23	46.07
2014	86.02	62.23	42.53
2015	85.38	61.38	41.80
2016	88.86	66.23	46.07
2017	88.86	66.23	46.07
2018	94.21	75.56	55.13
<b>Average:</b>	82.52	59.11	40.32
<b>Calculated:</b>	80	62	46

the proton density case, no statistically significant correlation was discovered. The last possibility for correlation remaining is to check if the proton velocity parameter was a small subset of days with both the density and velocity anomalies. Only 66 days (0.7 percent of the dataset) showed both density and velocity anomalies across the 23 year period. Only 12% (6/50) had an anomaly over the two-week period, 4% (2/50) during the one week period, and 2% (1/50) during the four-day period preceding the earthquake. There was no statistically significant correlation found.

### Increased Anomaly Threshold Test

Because of the skewness of the proton density distribution, bigger  $n$  standard deviation multipli-

ties could be used. The increase from five to nine standard deviations resulted in an average increase to 42 protons per cubic centimeter throughout the entire sample. This increase is 40% greater than the prior anomaly threshold, however, it has resulted in a 75% reduction in the number of anomalous days. An average of 11 days each year were marked as anomalous during the investigation period.

Only 28% (14/50) had an anomaly in the two-week period, 22% (11/50) in the one-week period, and 12% (6/50) in the four-day period before the earthquake when the entire sequence was repeated as previously shown for the proton density parameter. The predicted values were derived by recalculating the hypergeometric probability with the updated values (33 percent for the two-week pe-

riod, 18.67 percent for the one-week period, and 11 percent for the four-day period). As can be observed, the expected values derived using hypergeometric probability and the calculated values are within 10% of one another, which is enough to rule out any statistically significant correlation even with higher proton density values.

The proton velocity parameter was treated in the same way, with the  $n$  standard deviations increased from three to three and a half. This was a 6 percent increase, but it resulted in a 47 percent reduction in the number of anomalous days. In the two-week period, 28 percent (14/50) had an anomalous value, 14 percent (7/50) had an anomalous value in the one-week period, and 6 percent (3/50) had an anomalous value in the four-day period before an earthquake. There was no statistically significant correlation found.

It is worth noting that even when the dynamical anomaly threshold is disregarded (the threshold is based on the mean yearly proton density/velocity value and the standard deviation for the year) and an exceptionally high and constant anomaly threshold is applied ( $100 \text{ cm}^{-3}$  for proton density and  $1000 \text{ km/s}$  for proton velocity), only six individual anomalous days for the proton density are observed and 14 for the proton velocity parameter. Those two figures, respectively, represent 0.07 percent and 0.1 percent of the overall data set. Only one earthquake (Albania, April 9<sup>th</sup>, 2001) displayed a proton density anomaly prior to the earthquake, while none had a proton velocity anomaly. There was no statistically significant association, as with the other examples previously shown.

## Conclusion

A statistical correlation between increased solar wind parameters (proton density and velocity) and earthquakes on the Balkan Peninsula zone between 1996 and 2018 was presented in this study paper.

In the two weeks leading up to an earthquake, the presented increased proton density parameters show an 80% (40/50) correlation. This is the highest correlation found in this study. There is no substantial statistical correlation between these two occurrences on the Balkan Peninsula zone, according to statistical verification.

Proton velocity, like proton density, had no statistically significant correlation with earthquakes on the Balkan Peninsula, even when they occurred 56 percent of the time prior to earthquakes. A small group of days with both proton density and velocity anomalies produced no statistically significant correlation when studied.

Increasing the anomaly threshold for the proton density parameter from five to nine standard deviations and for the proton velocity parameter from three to three and a half standard deviations reduced the number of anomalous days by 75 and 47 percent, respectively. This reduction in the number of anomalous days did not result in a statistically meaningful improvement in the association. Furthermore, ignoring a dynamical anomaly threshold and concentrat-

ing on rare occurrences ( $100 \text{ cm}^{-3}$  and  $1000 \text{ km/s}$ ) did not produce any significant results.

In summarizing the subjectivities in this study, the primary focus must be on the selection of the research area. MARCHITELLI et al. (2020) found a positive association on a worldwide scale, however, this is of no practical utility because knowing when an earthquake will occur without knowing where it will occur is useless. The focus of this research article was on a smaller scale with a complex geological setting, however, the subjectivity involved in selecting such a location could not be adequately quantified.

The second subjectivity involves expanding the time span of 6.9 days stated by STRASER & CATALDI (2014). As the time window is extended, days that are more anomalous will surely occur before an earthquake. This subjectivity, like the preceding one, cannot be measured, but it is thought that statistical verification will eliminate it.

The last subjectivity is that when kurtosis increases, the value of variance decreases (TABACHNIK et al. 2007), and so the value of standard deviation decreases. If the standard deviation is less than the true value, the anomalous threshold will be less than the true value as well. Although this subjectivity is not quantifiable, it is eliminated by the use of the test of increasing anomalous thresholds.

Even when viewing this research topic with a critical mindset, it is believed that subjectivities in this study did not have a substantial, if any, influence. GELLER et al. (1997) also express skepticism regarding the discovery of a new, major precursor, stating that the probability of discovering the next one decreases with each new attempt, and that the possibility of discovering a novel precursor is now exceedingly minimal.

To locate a novel precursor candidate, you must first have a well-understood and well-established mechanism for such an occurrence. There is no documented and verified mechanism that could explain the solar wind generating earthquakes phenomenon, for example. The earthquakes are assumed to be tectonically controlled, but the increase in solar wind parameters acts as a trigger.

Even though the presented research found no statistically significant association for the Balkan Peninsula, it does not rule out the positive correlation found by MARCHITELLI et al. (2020) on a worldwide

scale. To establish it as a viable method, first, the mechanism should be established, then the usage of a statistical method to demonstrate the statistical significance of such a method with high confidences and low false alarm values should be undertaken. This approach has yet to be established for any of the known precursors.

## References

- CATALDI, G., CATALDI, D. & STRASER, V. 2016. Solar activity correlated to the M7.0 Japan earthquake occurred on april 15, 2016. *New Concepts in Global Tectonics Journal*, 4 (2): 279–285.
- CHARGE, ELEMENT, AND ISOTOPE ANALYSIS SYSTEM, Retrieved from <https://11.umd.edu/>, [1. 12. 2020]
- CICERONE, R.D., EBEL, J.E. & BRITTON, J. 2009. A systematic compilation of earthquake precursors. *Tectonophysics*. 476 (3–4): 371–396.
- GELLER, R.J., JACKSON, D.D., KAGAN, Y.Y. & MULARGIA, F. 1997. Earthquakes cannot be predicted. *Science*. 275(5306): 616–1616.
- GOOGLE EARTH, Retrieved from <https://earth.google.com/web/>, [7. 4. 2021]
- KAMER, Y., NANDAN, S., OUILLO, G., HIEMER, S. & SORNETTE, D. 2021. Democratizing earthquake predictability research: introducing the RichterX platform. *The European Physical Journal Special Topics*, 230 (1): 451–471.
- MARCHITELLI, V., HARABAGLIA, P., TROISE, C. & DE NATALE, G. 2020. On the correlation between solar activity and large earthquakes worldwide. *Scientific reports*, 10 (1): 1–10.
- MULARGIA, F. 1997. Retrospective validation of the time association of precursors. *Geophysical Journal International*, 131 (3): 500–504.
- MULARGIA, F. 2001. Retrospective selection bias (or the benefit of hindsight). *Geophysical Journal International*. 146 (2): 489–496.
- RHOADES, D.A. & EVISON, F.F. 1989. On the reliability of precursors. *Physics of the Earth and Planetary interiors*, 58 (2–3): 137–140.
- NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION, Retrieved from <https://www.ngdc.noaa.gov/mgg/topo/gltiles.html>, [22. 2. 2021]
- SCHOBER, P., BOER, C. & SCHWARTE, L.A. 2018. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, 126 (5): 1763–1768.
- SILSO DATA, ROYAL OBSERVATORY OF BELGIUM, BRUSSELS, Retrieved from <http://www.sidc.be/silso/infosnmtot>, [26. 2. 2021]
- SIMPSON, J.F. 1967. Solar activity as a triggering mechanism for earthquakes. *Earth and Planetary Science Letters*, 3: 417–425.
- STRASER, V. & CATALDI, G. 2014. *Solar wind proton density increase and geomagnetic background anomalies before strong M6+ earthquakes*. Space Research Institute of Moscow, Russian Academy of Sciences, MSS-14.
- TABACHNICK, B.G. FIDELL, L.S. & ULLMAN, J.B., 2007. *Using multivariate statistics Vol. 5*. Boston, MA: Pearson.
- UNITED STATES GEOLOGICAL SURVEY (USGS), Retrieved from <https://earthquake.usgs.gov/earthquakes/search/>, [13. 11. 2020]

## Резиме

### Корелабилност соларног ветра са сеизмичким догађајима у зони Балканског полуострва

Термин прекурсор потреса описује велики број различитих физичких феномена који се користе за предвиђање потреса (CICERONE et al., 2009). Предвиђање потреса представља правац сеизмологије који се бави краткорочним, средњерочним и дугорочним прогнозирањем потреса. Детерминистички приступ предвиђању потреса захтева познавање географске ширине и дужине епицентра потреса, магнитуду и време потреса, што се сматра нереалним за разлику од пробабилистичног приступа (врло слично временској прогнози), који даје вероватноћу да ће се догодити потрес за дати регион у одређеном временском периоду. Исто тако, свака метода предвиђања потреса треба да буде искључиво базирана на статистици, односно вероватноћи (KAMER et al., 2021).

Како би се предвиђање потреса извело са што већом сигурношћу, потрага за новим прекурсорима је веома актуелна. Данас, једну занимљиву групу прекурсора чине астрономски прекурсорни.

У раду је приказана статистичка корелација између повећаних вредности параметара соларног ветра (густина и брзина протона) и појаве потреса у региону Балканског полуострва у периоду од 1996. до 2018. године.

Приказани резултати за параметар густине протона показују да је у периоду од 14 дана пред потрес аномалија густине протона била је при-

сутна у 80 % случајева (40/50). Оваква вредност корелације је уједно и највећа добијена корелација у истраживању. Статистичка верификација значајности добијених резултата показала је да не постоји никаква значајна статистичка корелација између повећане вредности густине протона и потреса у региону Балканског полуострва.

Приказани резултати за параметар брзине протона највећу корелацију показују у вредности од 56 % (28/50) за референтни период од 14 дана пред потрес. Као и за параметар густине протона, статистичка верификација добијених резултата није показала значајну повезаност између две појаве.

Вршена је провера мањег подкупа аномалних појава соларног ветра, тј. појава аномалије густине и аномалије брзине протона истовремено (у истом дану). За цео истражни сет података постоји укупно 66 таквих дана, односно такви дани чине само 0.7 % истражног сета података. Корелација такве појаве са потресима приказала је да 12 % (6/50) потреса има аномалију и густине и брзине протона у референтном периоду од 14 дана. Статистичка провера овакве корелације није прешла постављени праг значајности појаве од 10 %, односно статистички значајна корелација није пронађена.

Пошто приказане појаве нису показале статистичку значајност са границама аномалије у вредности од пет и три стандардне девијације, вршено је повећање вредности границе аномалије ради провере статистичке корелације са мање аномалних дана. Иако се број аномалних дана за параметар густине протона смањило за око 75 %, а за параметар брзине протона се број аномалних дана смањило за 47 %, статистички значајна корелација није пронађена. Даљим занемаривањем динамичке границе аномалије (зависност границе аномалије од средње вредности и стандардне девијације за дату годину) и постављањем константне границе за аномалије за цео истражни сет података од  $100 \text{ cm}^{-3}$  за густину протона и  $1000 \text{ km/s}$  за брзину протона, такође није пронађена значајна статистичка корелација.

Највећа субјективност у овом истраживању своди се, пре свега, на одабир истражног простора, односно на ослањање на тектонске карактери-

стике. Пошто је приказана корелација у MARTICHELLI et al. (2020) на глобалном величинском подручју, те било каква спознаја да ће се потрес догодити нема практичну примену, јер није могуће тачно одредити географску ширину и дужину епицентра потреса, сматрало се да је потребно смањити истражно подручје ради ужег одређивања епицентра потреса. Сматра се да није могуће квантификовати овако унешену субјективност.

Друга значајна субјективност представља повећање референтног периода пред потрес у односу на пријављени референтни период према STRASER & CATALDI (2014). Повећавањем референтног периода повећава се и број потреса који има аномалију. Као и претходна субјективност, сматра се да се ни ова субјективност не може квантификовати.

Последњу субјективност у овом истраживању представља чињеница да повећане вредности куртозе смањују вредност варијанце (ТАВАСНИК et al. 2007), а тиме утичу и на смањење вредности стандардне девијације. Ако су вредности стандардне девијације умањене у односу на њихову истиниту вредност, онда је и граница аномалије умањена за исту ту вредност. Овакву субјективност није потребно квантификовати, већ се повећавањем границе аномалије и одређивањем статистичке значајности такве појаве она и уклонила.

Иако се овом истраживању приступило са скептичног аспекта проналаска позитивне и статистички значајне корелације између повећаних вредности параметара соларног ветра и потреса, сматра се да су у овом истраживању субјективности максимално смањење. GELLER et al. (1997) такође приказује скептичан приступ проналаска поузданог прекурсора потреса, где се наводи да се сваким новим покушајем *a priori* вероватноћа проналаска наредног прекурсора смањује, као и да је тренутна вероватноћа проналаска поузданог прекурсора екстремно мала.

Да би одређена појава била добар кандидат за прекурсор потребно је најпре имати утемељен механизам који доводи до потреса (или бар убрзава потрес). И даље није познат, нити квантификован утицај соларног ветра на потресе, сматра да су потреси и даље тектонски кон-

тролисани, а само евентуално „убрзани“ повећаним вредностима параметара соларног ветра, ако такав механизам постоји.

Приказано истраживање, иако је показало мањак статистичке корелације повећаних вредности параметара соларног ветра и потреса на Балканском полуострву, није искључило постојање позитивне корелације на глобалном величинском подручју. Да би се прецизно одредио утицај соларног ветра на потресе потребно је прво одредити механизам утицаја који ће бити

одређен мереним подацима. Након тога, потребно је одредити статистички значајну методу која предвиђа географску ширину и дужину епицентра, као и време потреса, са високом поузданошћу и малим вредностима стопе лажних аларма. Оваква метода за сада није поуздано откривена ни за један пријављен прекурсор (или потенцијални узрочник).

*Manuscript received October 29, 2021*

*Revised manuscript accepted November 22, 2021*

