Investigation of Relation between Solar Activity and Earthquakes with Deep Learning Method

Leilei Li[1], Hong Gu[1], Ryosuke Kikuyama[2], Ryuei Nishii[3], Pan Qin[1]

[1]Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China
[2]Mazda Motor Company, Hiroshima, Japan
[3]School of Information and Data Sciences, Nagasaki University, Nagasaki, Japan
Email: qp112cn@dlut.edu.cn

Abstract

Solar activity (SA) has been hypothesized to be a trigger of earthquakes, although it is not as intuitively associated as other potential triggers such as tidal stress, rainfall, and the building of artificial water reservoirs. Here, we investigate the relation between SA and global earthquake numbers (GEN) by using a deep learning method to test the hypothesis. We use the daily data of GEN and SA (1996/01/01-2019/12/31) to construct a temporal convolution network (TCN). From the computational results, we confirm that the TCN captures the relation between SA and earthquakes with magnitudes from 4.0 to 4.9. We also find that the TCN achieves better fitting and prediction performance compared with previous work.

Keywords

Deep Learning, Earthquakes, Prediction, Solar Activities, Temporal Convolution Network

1. Introduction

The 11-year solar cycle contributes to events such as sunspots, coronal mass ejection, and solar wind. The mechanism of the sun-earth magnetosphere connection is a mystery in relation to earthquakes [1]. Several studies have proposed that solar activity (SA) might be linked to earthquakes [2] [3] [4]. Statistical methods are usually used to prove this hypothesis. Reference [5] suggests a correlation between SA and large earthquakes worldwide, and [6] investigates the correlation between long-range clustering of global seismicity and SA. Sunspot number is also considered to be an SA variable for predicting earthquakes [7]. Meanwhile, some mechanisms have been considered to improve the correlation between the SA and the

earthquakes. For example, induced current causes an increase in fault stress through piezoelectricity [8], and the eddy electric currents in faults reduce the shear strength [9].

The previous studies mainly focused on investigating the significant correlation between SA and earthquakes using non-parametric statistical methods. However, parametric statistical models and machine learning models are also necessary for earthquake forecasting, although they are far from applicable to this task. In our previous work [10], we attempted to predict Global Earthquake Numbers (GEN) by using variables associated with SA as inputs. The results in [10] show that the GEN of earthquakes with magnitude 4 - 4.9 is most predictable.

With the development of sensing technologies, including GPS and InSAR [11], a massive amount of data on SA has been accumulated. Furthermore, the solar-earth coupling can be characterized as a non-linear dynamical system. For these two reasons, we decided to construct deep learning models to predict GEN with SA as the input for earthquakes of magnitude 4 - 4.9. In particular, we considered daily time series of GEN and SA in sequential format. The recurrent neural network (RNN) and long-short-memory-term are two benchmark DL models for sequential data. However, feedback in the recurrent architecture can lead to higher computational complexity [12]. Recent studies [13] [14] indicate that certain convolutional neural network (CNN) architectures can reach state-of-the-art accuracy for sequential data. A CNN can ensure the causality of sequential data of any sequence length with no feedback.

By considering the proven effectiveness of CNNs for sequencing data, we took all the observations in the time-series format and implemented the temporal convolutional network (TCN) [15]. We constructed TCN by using GEN data and SA data as input to predict GEN for earthquakes of magnitude 4 - 4.9.

2. Dataset

Daily data of GEN were downloaded from ComCat (https://earthquake.usgs.gov/earthquakes/search/). The data ranged from 01/01/1996 to 12/31/2019, including the 23rd and 24th solar cycles, and are partly depicted in Table 1. *EQi* means earthquakes with magnitude i - i.9 for $i = 3, 4, 5, 6, 7$. Note that earthquakes with $M \geq 8$ rarely occurred, so we combined those into one column: *EQ*89. The data contain an earthquake $M = 7.2$ (04/05/2010) that occurred in Estado de Baja California of Mexico and the Touhoku earthquake $M = 9.0$ (03/11/2011) that occurred in the north-east of Japan. Because large earthquakes always cause aftershocks, the GEN itself was also used as an input of TCN. (Figure 1)

The daily data of SA were downloaded from OMNIWeb (https://omniweb.gsfc.nasa.gov/). The SA variables used in this research are listed in Table 2. Part of the SA data are illustrated in Figure 2. Missing values in the original SA data were filled using the linear interpolation method. (Table 3)

Table 1. Daily data of GEN.

Time	EQ3	EQ4	EQ5	EQ6	EQ7	EQ89
1996/1/1	69	54	10	2	1	0
1996/1/2	50	32	5	0	0	0
1996/1/3	25	32	4	0	0	0
1996/1/4	28	41	5	0	0	0
1996/1/5	14	21	1	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
2019/12/27	13	26	2	0	0	0
2019/12/28	13	32	5	0	0	0
2019/12/29	14	36	4	0	0	0
2019/12/30	12	34	4	0	0	0
2019/12/31	9	28	1	0	0	0
Sum	144,258	258,140	36,766	3275	339	26

Table 2. SA variables list.

Notations	SA variables	units
B	IMF Magnitude	nT
T	Proton temperature	B
D	Proton density	n/cc
V	Plasma (Flow) speed	km/s
P	Flow pressure	nPa
E_Y	Earth's electric field east-west component	mV/m
SSN	Sunspot number	n
DST	Disturbance storm time index	nT
PCI	Polar cap index	

Table 3. Daily data of SA variables.

| Time	B	T	D	V	P	E_Y	SSN	DST	PCI
1996/1/1	4.9	72,020	7.9	403	2.39	−0.42	0	−2	0.2
1996/1/2	4.9	77,660	8	398	2.45	0.23	14	−6	0.6
1996/1/3	5.9	132,150	5.2	485	2.39	−0.26	22	−8	0.6
1996/1/4	4.4	93,895	5.5	443	2.06	0.01	35	−5	0.3
1996/1/5	3.9	50,191	10.2	416	3.12	0.08	56	−1	0.6
...	...	...	...	...	...	...	...	...	...
2019/12/27	3.9	53,858	5.9	371	1.41	−0.01	0	−2	0.3
2019/12/28	3.4	38,546	6.4	345	1.33	−0.03	0	−1	0.1
z2019/12/29	4.1	22,298	9.1	309	1.47	−0.01	0	−1	0.1
2019/12/30	5.6	20,010	13.3	298	2.02	−0.35	0	5	0.1
2019/12/31	4.3	25,823	13.7	306	2.23	0.38	0	1	0.4
mean	5.76	93,432.44	6.22	428.96	2.01	0.02	65.38	−12.16	0.95
std	2.56	72,500.54	3.89	96.86	1.21	0.76	63.92	17.56	0.85
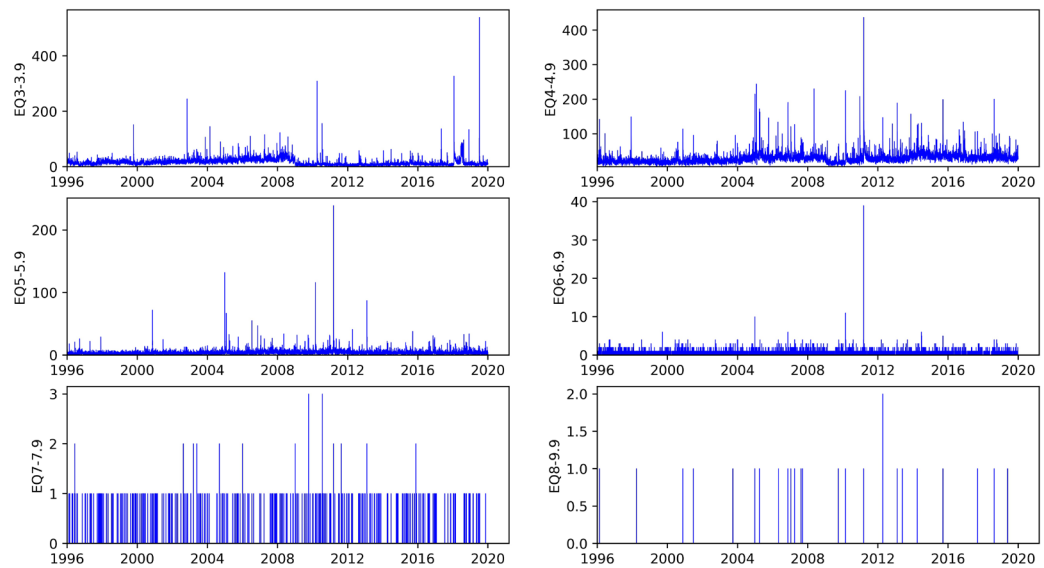
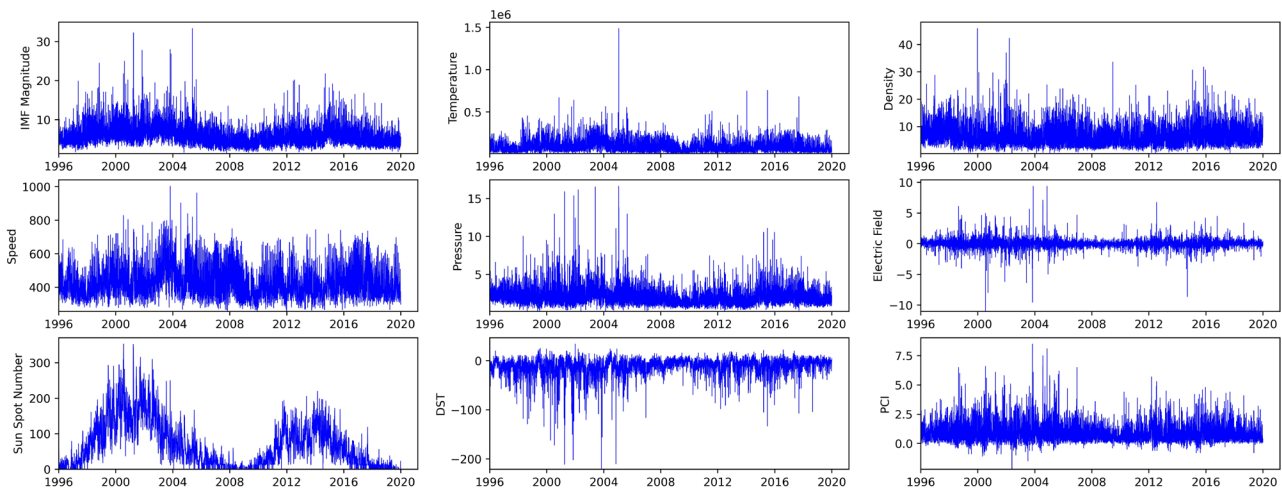**Figure 1.** Time series of earthquake numbers (1996/01/01-2019/12/31).

Figure 2. Time series of SA variables (1996/01/01-2019/12/31).

3. TCN Architecture

According to our previous works [10], to model the earthquakes, GEN of *EQ*4 is taken as the outputs of TCN, denoted as $y(k)$. Then, two types of inputs for TCN are included: the first type, denoted as $x_1(k)$, takes GEN data up to k as the inputs of TCN; the second type, denoted as $x_2(k)$, includes SA variables. Here, k means that the latest observations in $x_1(k)$ and $x_2(k)$ are obtained at time k. With respect to our previous results [10], the maximum time lag of each variable was set as 14 days in both $x_1(k)$ and $x_2(k)$. That is,

$$x_1(k) \equiv \left[EQ3(k), \cdots, EQ3(k-13), \cdots, EQ89(k), \cdots, EQ89(k-13) \right]^{\mathrm{T}}$$

$$x_2(k) \equiv \left[B(k), \cdots, B(k-13), \cdots, PCI(k), \cdots, PCI(k-13) \right]^{\mathrm{T}}$$

In this way, we construct a non-linear model

$$y(k+d) = g\big(\pmb{x}_1(k), \pmb{x}_2(k)\big) + \varepsilon(k)$$

to couple the relation between GEN and SA by comparing these models, with $\varepsilon(k)$ being an independently identically distributed Gaussian noise term. Here, d is the number of days later than day k and indicates the prediction steps. In this research, $k = 1, 2, 3$ is considered.

This research uses TCN as $g(\cdot)$, whose architecture is shown by **Figure 3**. The TCN is mainly composed of convolutional blocks with 16, 32, 32, and 64 channels. In each block, a dilated convolutional operation is performed on a sequence input $\pmb{u} \in \mathbb{R}^n$:

$$F(s) = (u_{*d} f)(s) = \sum_{i=0}^{m-1} f(i) \cdot u_{s-d}$$

with $f : \{0, 1, \cdots, m-1\} \to \mathbb{R}$ being a filter.

Because the maximum time lag is relatively short, the convolutional kernels of size 1×2 are implemented in each block. To obtain the robust estimates, the Huber loss function is used as follows:
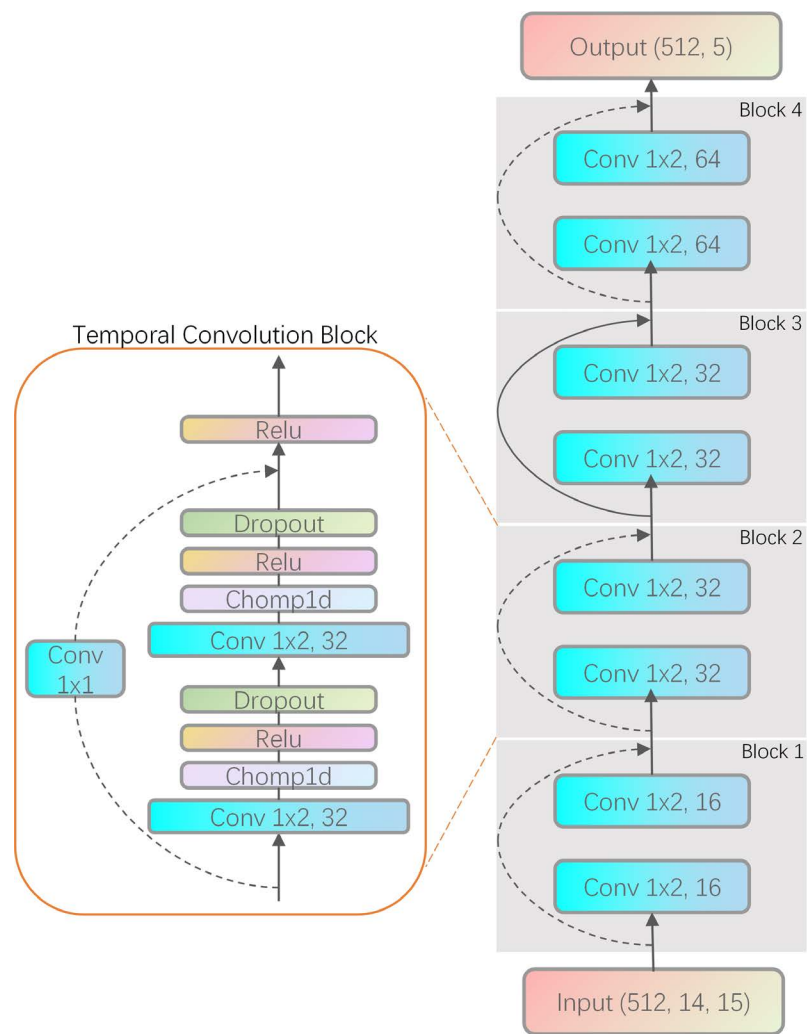

Figure 3. Architecture of TCN.

$$L_\delta\left(y_i, f\left(x_i\right)\right) = \begin{cases} \dfrac{1}{2}\left(y_i - f\left(x_i\right)\right)^2, & \text{if } \left|y_i - f\left(x_i\right)\right| \le \delta \\ \delta\left|y_i - f\left(x_i\right)\right| - \dfrac{1}{2}\delta^2, & \text{otherwise} \end{cases}$$

where y_i is the observed GEN at time i, and x_i and $f\left(x_i\right)$ are the corresponding input and output of TCN. The Adam optimizer [16] was used to chain TCN. The SA and GEN data in the 23rd solar cycle (01/01/1996-12/31/2007) were used as the training data. The SA and GEN data in the 24th solar cycle (01/01/2008-12/31/2019) were used as the test data.

4. Prediction Results

The whole dataset was divided into two parts. The SA and GEN data in the 23rd solar cycle (01/01/1996-12/31/2007) were used as the training data. The SA and GEN data in the 24th solar cycle (01/01/2008-12/31/2019) were used as the test data to verify the trained TCN. Pearson's correlation coefficient R was used to evaluate the fitting and prediction performance of the TCN.

4.1. TCN without/with SA Variables

First, we constructed a TCN without SA variables. **Figure 4** illustrates the loss curves of the training and test losses versus epoch number. The curves indicate that 100 epochs are enough to ensure convergence of the TCN training.

We also constructed TCN with all of the SA variables. **Figure 5** illustrates the training and test losses plotted against epoch number. The curves indicate that 100 epochs are enough to ensure convergence of the TCN training.

Table 4 lists the fitting and prediction performance of TCNs without SA variables for 1- to 3-day-ahead predictions. Let R_f and R_p be the correlations between the real observations of *EQ*4 and the output of the TCN obtained from the training data and test data, respectively. **Table 5** lists the fitting and prediction
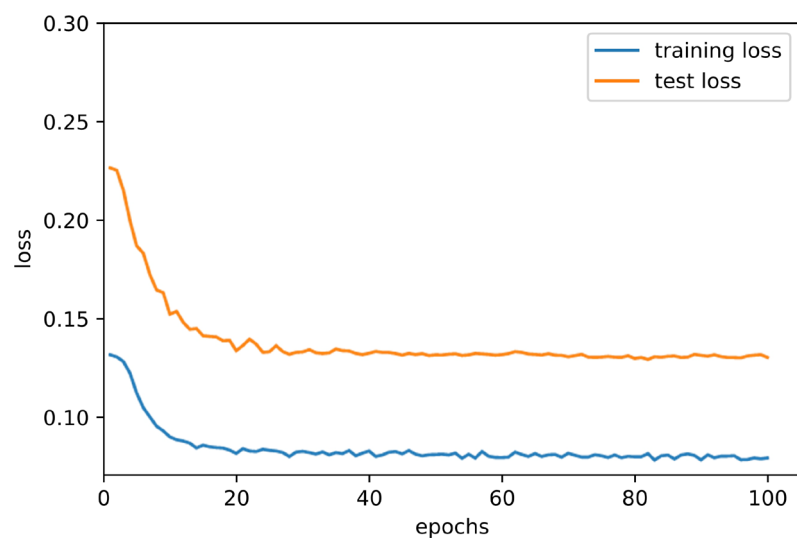

Figure 4. Dynamical curves of traning and test losses versus epoch number for TCN without SA variables.
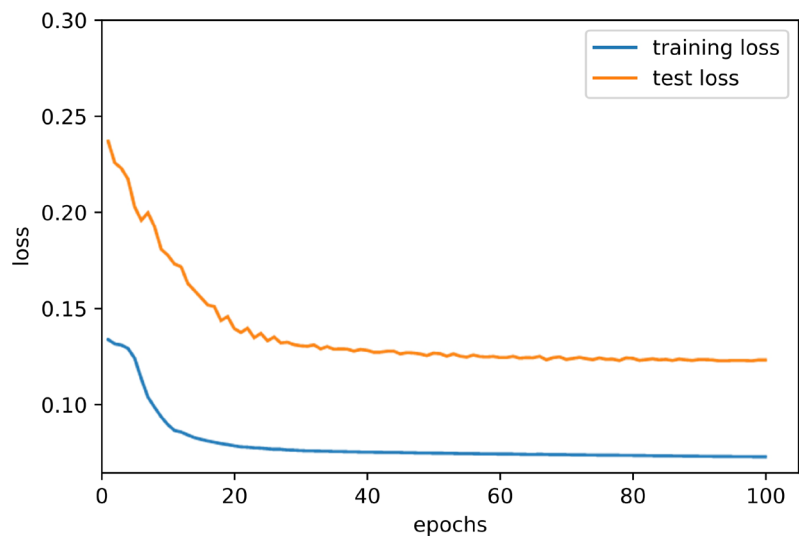
Figure 5. Dynamical curves of traning and test losses versus epoch number for TCN with SA variables.

Table 4. Fitting and prediction performance of TCNs without SA variables.

d	R_f	R_p	decrease
1	0.6702	0.6222	0.0480
2	0.5115	0.4040	0.1075
3	0.4856	0.3240	0.1616

Table 5. Fitting and prediction performance of TCNs with SA variables.

d	R_f	R_p	decrease
1	0.7131	0.6502	0.0629
2	0.5324	0.4384	0.0940
3	0.5088	0.3496	0.1592

performance of TCNs with SA variables for 1- to 3-day-ahead predictions. As a reasonable result, R_f is R_p for all days ahead in **Table 4** and **Table 5**. Thus, the "decrease" in **Table 4** and **Table 5** means the difference between R_f and R_p.

The two tables indicate that the TCNs are of better fitting and prediction performance than the support vector regression in our previous work [10]. By comparing **Table 4** and **Table 5**, it can be seen that the SA variables improve both the fitting and prediction performance of TCNs. The gap between R_f and R_p is trivial in the 1-day-ahead prediction, which supposes a balance between the fitting and prediction performance of TCNs with/without SA variables. However, R_p significantly decreases for the 2- and 3-day-ahead predictions.

4.2. Impact of SA Variables on Prediction of Earthquakes

To evaluate how the SA variables improve the prediction of earthquakes, we adopted the following forward stepwise procedure:

Table 6. Variables selected by forward stepwise procedure.

Step	Variable	R_f	R_p	R_a
1	*EQ*4	0.6352	0.6098	0.6225
2	*EQ*5	0.6701	0.6296	0.6499
3	*V*	0.6844	0.6390	0.6617
4	*EQ*89	0.6798	0.6526	0.6662
5	*EQ*6	0.6734	0.6500	0.6618
6	*SSN*	0.6736	0.6485	0.6610
7	E_y	0.6846	0.6447	0.6647
8	*EQ*3	0.6905	0.6365	0.6635
9	*D*	0.6833	0.6472	0.6653
10	*P*	0.6962	0.6331	0.6646
11	*EQ*7	0.6912	0.6461	0.6687
12	*T*	0.7064	0.6445	0.6755
13	*PCI*	0.6987	0.6425	0.6705
14	*DST*	0.6931	0.6387	0.6659
15	*B*	0.7131	0.6502	0.6816

1) Let $\mathcal{V}$ denote the set of inputs of TCN, which is initially assumed to be the empty set. Let
$\mathcal{C} = \{EQ3, EQ4, EQ5, EQ6, EQ7, EQ89, B, T, D, V, P, E_y, SSN, DST, PCI\}$ be the set of 15 input candidates of the TCN.

2) Add one variable from $\mathcal{C}$ to $\mathcal{V}$ and construct a TCN. Define the corresponding evaluation criterion $R_a = (R_p + R_f)/2$. Add to $\mathcal{V}$ the variable from $\mathcal{C}$ that gives the biggest improvement in R_a.

3) Repeat (2) until $\mathcal{C}$ becomes empty and a total of 15 TCNs are obtained.

Table 6 shows the sequentially selected variables according to R_a for the 1-day-ahead prediction. We can see that the plasma speed *V* improves R_a by almost 0.02 based on *EQ*3 and *EQ*4 in step 3. The IMF Magnitude improves R_a by almost 0.02 at the last step, jointly with other variables. These results suggest that all the SA variables should be used as the inputs of TCNs.

5. Conclusions

In this research, we investigate the relation between SA and GEN. We construct the deep learning model TCN to predict *EQ*4 for 1- to 3-day-ahead predictions. The numerical results show that:

1) Compared with SVR in our previous works, TCN significantly enhances the fitting and prediction performance. This result confirms that there exists a strong nonlinear relation between GEN and SA.

2) Because the fitting performance R_f is similar to R_p, we suppose that TCN is of potential capacity for the 1-day-ahead prediction for *EQ*4.

3) *EQ*4 in the past is the crucial input of TCN. Thus, TCN is essentially a non-linear autoregressive model. However, SA variables can still improve the fitting and prediction performance of TCN.

From the aforementioned results, we suppose that SA has the potential to affect GEN.

TCNs in this research are still far from being predictive. Table 6 shows that the TCN is continuously improved until all the SA variables are implemented. This result suggests that the prediction performance can be further improved by considering more variables other than the candidates selected in this research. Over the decades, lots of novel geophysics and space data have become available, thanks to improvements in sensing and measurement technologies. Although earthquakes remain not predictable for now, we will continue to reveal relations among earthquakes, the earth's environment and SA on the basis of various statistical methods and machine/deep learning models.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Sukma, I. and Abidin, Z.Z. (2017) Study of Seismic Activity during the Ascending and Descending Phases of Solar Activity. *Indian Journal of Physics*, **91**, 595-606. https://doi.org/10.1007/s12648-016-0943-5

[2] Odintsov, S., Boyarchuk, K., Georgieva, K., Kirov, B. and Atanasov, D. (2006) Long-Period Trends in Global Seismic and Geomagnetic Activity and Their Relation to Solar Activity. *Physics and Chemistry of the Earth*, *Parts A/B/C*, **31**, 88-93. https://doi.org/10.1016/j.pce.2005.03.004

[3] Simpson, J.F. (1967) Solar Activity as a Triggering Mechanism for Earthquakes. *Earth and Planetary Science Letters*, **3**, 417-425. https://doi.org/10.1016/0012-821X(67)90071-4

[4] Huzaimy, J.M. and Yumoto, K. (2011) Possible Correlation between Solar Activity and Global Seismicity. *Proceeding of the* 2011 *IEEE International Conference on Space Science and Communication*, Penang, 12-13 July 2011, 138-141. https://doi.org/10.1109/IConSpace.2011.6015869

[5] Marchitelli, V., Harabaglia, P., Troise, C. and De Natale, G. (2020) On the Correlation between Solar Activity and Large Earthquakes Worldwide. *Scientific Reports*, **10**, Article No. 11495. https://doi.org/10.1038/s41598-020-67860-3

[6] Marchitelli, V., Troise, C., Harabaglia, P., Valenzano, B. and De Natale, G. (2020) On the Long Range Clustering of Global Seismicity and Its Correlation with Solar Activity: A New Perspective for Earthquake Forecasting. *Frontiers in Earth Science*, **8**, 470. https://doi.org/10.3389/feart.2020.595209

[7] Khain, V.E. and Khalilov, E.N. (2007) About Possible Influence of Solar Activity upon Seismic and Volcanic Activities: Long-Term Forecast. *Transactions of the International Academy of Science H&E*, **3**, 217-240.

[8] Love, J.J. and Thomas, J.N. (2013) Insignificant Solar-Terrestrial Triggering of Earthquakes. *Geophysical Research Letters*, **40**, 1165-1170.

https://doi.org/10.1002/grl.50211

[9]   Han, Y., Guo, Z., Wu, J. and Ma, L. (2004) Possible Triggering of Solar Activity to Big Earthquakes (Ms ≥ 8) in Faults with near West-East Strike in China. *Science in China Series G: Physics and Astronomy*, **47**, 173-181. https://doi.org/10.1360/03yw0103

[10] Nishii, R., Qin, P. and Kikuyama, R. (2020) Solar Activity is One of Triggers of Earthquakes with Magnitudes Less than 6. *IGARSS* 2020 *IEEE International Geoscience and Remote Sensing Symposium*, Hawaii, 26 September-2 October 2020, 377-380. https://doi.org/10.1109/IGARSS39084.2020.9323381

[11] Qiao, X., Wang, Q., Yang, S., Li, J., Zou, R. and Ding, K. (2015) The 2008 Nura Mw6.7 Earthquake: A Shallow Rupture on the Main Pamir Thrust Revealed by GPS and InSAR. *Geodesy and Geodynamics*, **6**, 91-100. https://doi.org/10.1016/j.geog.2015.01.005

[12] Goodfellow, I., Bengio, Y. and Courville, A. (2016) Deep Learning. MIT Press, Cambridge.

[13] Dauphin, Y.N., Fan, A., Auli, M. and Grangier, D. (2017). Language Modeling with Gated Convolutional Networks. *Proceedings of the* 34*th International Conference on Machine Learning*, Sydney, 6-11 August 2017, 933-941.

[14] Oord, A.V.D., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A. and Kavukcuoglu, K. (2016) Wavenet: A Generative Model for Raw Audio. https://ui.adsabs.harvard.edu/abs/2016arXiv160903499V

[15] Bai, S., Kolter, J.Z. and Koltun, V. (2018) An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. https://ui.adsabs.harvard.edu/abs/2018arXiv180301271B

[16] Kingma, D.P. and Ba, J. (2014) Adam: A Method for Stochastic Optimization. https://ui.adsabs.harvard.edu/abs/2014arXiv1412.6980K