4. *Оливерос А.П., Хлебосолова О.А.* Ландшафты зоологического заповедника Дельта дель Кауто (Куба) и их изучение // IV Международная научно-практическая конференция «Добродеевские чтения 2020», МГОУ. С. 213-218.

5. DECRETO - LEY NO. 201 del Sistema Nacional de Área Protegidas. Gaceta Oficial de la República de Cuba. Número 84. 1999. Pp. 1355-1356.

6. *Hernández E.* La implementación del Plan Estratégico del Sistema Nacional de Áreas Protegidas de Cuba // Cub@: Medio Ambiente y Desarrollo. No.13, 2007 pp 7.

7. Informe SNAP 2018. Habana 2018. pp 18.

8. *Mateo, J.* Planificación Ambiental. Facultad de Geografía de la Universidad de la Habana. La Habana, Cuba: Editorial Félix Varela. 2008. pp 166.

9. *Navarro, S.* El estudio de los índices de vegetación como base para conocer las relaciones entre la vegetación y el clima. 2010, pp. 1095-1108.

10. *Oliveros A.P., Khlebosolova O.A.* Comparative analysis of the spectral response of vegetation for monitoring the state of the landscapes of the protected area of Delta del Cauto (Cuba) / Mediterranean Geosciences Union Annual Meeting (MedGU-21), 25–28 November 2021, Istanbul, Turkey. P. 281.

11. *Oltremari, J.V. y Thelen, Kyran D.* (2003): "Evolution of the Planning Process for Protected Areas in Latin America". Natural Areas Journal. Santiago de Chile. FAO-RLC. Volume 23 (2). 174–179 pp.

12. Tabloide de curso de Áreas Protegidas de Cuba y Patrimonio Nacional. Editorial Academia. La Habana 2003. pp 31.

A.P. Oliveros, O.A. Khlebosolova

IMPROVING CUBA'S PROTECTED AREA SYSTEM FOR THE CONSERVATION OF RARE SPECIES, LANDSCAPES AND EFFECTIVE MANAGEMENT

Abstract: The article discusses the history and current state of the Cuba's protected area system. On the example of the zoological reserve Delta del Cauto, a program of landscape and ecological research was proposed and tested to preserve biological diversity, rare species of flora and fauna, and natural landscapes within protected natural areas.

Key words: protected natural areas, Cuba, professional competence, program of landscape and ecological research, biological diversity, rare species.

M.M. Radovanović[a,b], S. Malinović-Milićević[a], S. Radenković[c], M.Milenković[a], B.Milovanović[a], A. Milanović Pešić[a], V.Popović[a]

[a]Geographical Institute "Jovan Cvijić" SASA, 9 Djure Jakšića St., 11000 Belgrade, Serbia *(E-mail: m.radovanovic@gi.sanu.ac.rs; s.malinovic-milicevic@gi.sanu.ac.rs; m.milenkovic@gi.sanu.ac.rs; b.milovanovic@gi.sanu.ac.rs; a.milanovic@gi.sanu.ac.rs; v.popovic@gi.sanu.ac.rs)*

[b]*Institute of Sports, Tourism and Service, South Ural State University, 76 Lenin A, 454080 Chelyabinsk, Russia*

[c]*Faculty of Banking, Insurance, and Finance, Belgrade Banking Academy, 12 Zmaj Jovina St., 11000 Belgrade, Serbia (E-mail: sonja.radenkovic@bba.edu.rs)*

INFLUENCE OF SPACE WEATHER ON PRECIPITATION NDUCED FLOODS –APPLYING OF SOLAR ACTIVITY TIME SERIES IN THE PREDICTION OF PRECIPITATION-INDUCED FLOODS BY USING THE MACHINE LEARNING

1. Introduction

Solar energy reaching the Earth varies on a wide range of time scales and correlates with atmospheric parameters [1]. The effect on Earth depends on solar activity intensity, the heliographic position of the eruptions, electromagnetic characteristics of the solar wind, as well as the conditions in interplanetary space [2,3]. However, it is difficult to estimate to which extent will influence the environmental processes and climate. Although the general acceptance of the existence of a link between the solar wind and the troposphere has not been achieved, a series of papers investigating empirical relationships between solar activity and climate variables have indicated the impact of the sun on climate. Veretenenko and Thejll [4,5] revealed that Solar Proton Events, with energies above 90 MeV are accompanied by the intensification of cyclonic activity at middle latitudes. According to [6] during solar proton penetration into Earth's atmosphere precipitation intensity in the territory of the former USSR increased by ~10%. Prikril et al. [7] revealed that heavy precipitation causing floods and flash floods tend to occur within several days of high-speed solar winds coming from the coronal hole. In [8] it is shown that slow solar wind becomes stronger when the coronal hole or active region approaches the geo-efficient position, while its effects on the Earth can be expected in 2-3 days. According to [9] the maximum amount of precipitation occurs 14 days after the solar wind is observed.

In this paper, we focused on establishing hidden dependencies between precipitation-induced floods in the United Kingdom (UK) and the flows of particles from the Sun based on 20 flood events in the UK from October 2001 to December 2019. To justify and also quantify the relationship we are advocating we used Machine Learning Classification Predictive Modelling.

2. Materials and Methods

2.1. Data description

To test for a possible connection between precipitation-induced floods and solar activity, we used several datasets and data sources. The output data were precipitations in the flooded area in the period 10 days before and during each flood event. The flood data used in this study were taken from the Emergency Events Database Center for Research on the Epidemiology of Disasters (EM-DAT database) [10]. For this analysis, 20 floods in the UK recorded in the EM-DAT database were selected from the period 2001-2019. Of the 20 selected floods, 17 are of rain origin, and the remaining 3 are not defined in detail. Input data were integral proton flux (p/cs2-sec-ster), differential electron and proton flux (p/cs2-sec-ster), solar wind characteristics (proton density (particles/cc), bulk speed (km/s), ion temperatures (degrees K)), and daily solar F10.7 cm radio flux. The 5-minute data on integral proton flux and differential electron and proton flux was provided by the Advanced Composition Explorer (ACE) Satellite [11]. Ranges for differential electron flux were 38-53 keV, and 175-315 keV, while for differential proton flux were 47-65 keV, 47-68 keV, 65-112 keV, 112-187 keV, 115-195 keV, 310-580 keV, 761-1220 keV, 795-1193 keV, 1060-1900 keV, and 1060-1910 keV. The data source for solar wind plasma data (3 per day) and 10.7 cm radio flux (1 per day) was the Laboratory for Atmospheric and Space Physics [12] at the University of Colorado Boulder (CU) for the period 2001-2004, and Space Weather Canada [13] for the period 2007-2019.

2.2 The impact of precipitation on the occurrence of floods

To detect the influence of the amount of precipitation on the occurrence of floods, the design of superimposed epochs was used [14]. To examine whether and on which days there is a significant influence of the amount of precipitation on the occurrence of floods, a one-way analysis of variance was applied and presented in Table 1. From Table 1 it can be concluded that there is a statistically significant difference between the amounts of precipitation during the observed days.

Table 1. *The results of a one-way analysis of the variance of the amount of precipitation on the day when the flood occurred and during the ten days preceding it*

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5018.900	10	501.890	11.665	0.000
Within Groups	9465.810	220	43.026		
Total	14484.710	230			

To determine on which day the precipitation was significantly different from the precipitation on each of the remaining days, Hochberg and Games-Howell posthoc tests were applied. These tests were chosen considering that Levene's test showed that there is no equality of variances for the analyzed variables. The Hochberg and Games-Howell tests revealed that the amount of precipitation on the day of the flood, as well as on the day preceding it, plays a significant role in the occurrence of floods. Similar findings for the UK were reached by [15,16].

2.2 Preliminary processing of input data and correlation analysis

To obtain the final data set for each flood event, we firstly performed binary classification on data. The positions of the peaks for the solar activity fields and precipitations were determined programmatically with subsequent manual verification. Then we made a correlation analysis between input factors and precipitations to find independent functional dependencies for each flood event separately. The results are shown in Table 2.

2.3 Machine learning and forecast of precipitation

Machine Learning Classification Predictive Modelling is applied to try to establish an eventual link between input and output data. In this study, we considered using the decision tree to allow clear understanding and justify the classification decision. The decision tree algorithm builds tree branches in a hierarchy approach. Each branch uses the if-then rule and divides the data set into subsets based on the most important characteristics. The main idea of the decision tree is to identify the most important characteristics that contain the most information about the target feature, and then divide the data set along with the values of these characteristics so that the values of the target characteristics in the resulting nodes are as clean as possible.

Table 2. *Maximum values of correlation coefficients between input factors and precipitations**

	integral proton flux		differential electron flux		differential proton flux										solar wind			radio flux
flood events	$> 10\ MeV$	$> 30\ MeV$	**38 – 53 keV**	**175 – 31 keV**	**47 – 65 keV**	**47 – 68 keV**	**65 – 112 keV**	**112 – 187 keV**	**115 – 195 keV**	**310 – 580 keV**	**761 – 1220 keV**	**795 – 1193 keV**	**1060 – 1910 keV**	**1060 – 1900 keV**	**BULK SPEED**	**ION TEMPERA-TURE**	**PROTON DENSITY**	**10.7 cm Radio Flux**
2001_0645	0.23	0.23	0.18	0.18	0.68	-	-	0.68	-	0.68	0.68	-	0.68	-	0.41	0.41	0.64	0.18
2002_0463	0.19	0.52	0.19	0.52	-	-	0.19	0.29	-	0.29	0.19	-	0.19	-	0.29	/	0.29	0.24
2002_0488	/	/	0.43	0.43	-	-	/	0.19	-	0.19	0.18	-	0.43	-	0.43	0.82	/	0.19
2002_0774	0.15	0.13	0.15	0.1	-	0.15	-	-	0.15	-	0.13	-	0.13	-	0.15	0.15	0.16	0.13
2004_0423	0.26	0.33	0.33	0.26	-	/	-	-	0.33	0.11	0.26	-	0.17	-	0.77	0.33	0.26	0.17
2007_0201	0.26	0.3	0.21	/	-	0.34	-	-	0.26	0.21	0.11	-	0.26	-	0.21	0.38	0.47	0.17
2007_0247	/	0.15	0.41	0.27	-	0.42	-	-	0.27	0.42	0.22	-	0.62	-	/	0.22	0.15	0.23
2007_0278	0.45	0.65	/	0.21	-	0.21	-	-	0.11	0.21	/	-	0.45	-	/	0.43	/	0.44
2008_0055	/	0.1	0.27	0.22	-	0.24	-	-	0.27	0.22	0.42	-	0.3	-	0.22	0.16	0.19	/
2008_0381	0.27	/	0.3	/	-	0.13	-	-	0.53	0.27	/	-	0.82	-	0.3	0.3	0.13	0.27
2009_0497	0.34	/	0.33	0.33	-	0.44	-	-	0.34	0.44	0.34	-	0.44	-	0.58	0.15	0.15	/
2012_0446	0.35	0.35	0.25	0.3	-	0.25	-	-	0.17	0.29	-	/	-	/	0.17	0.29	/	0.2
2012_0488	0.24	0.24	0.13	0.13	-	0.81	-	-	0.39	0.36	-	0.46	-	0.46	/	/	0.46	0.36
2012_0548	0.41	0.13	/	/	-	/	-	-	/	/	-	/	-	0.41	0.21	0.16	0.41	0.16
2012_0549	0.26	0.26	0.26	0.4	-	0.26	-	-	0.4	0.4	-	0.4	-	0.67	0.26	0.4	0.26	0.13
2012_0552	0.2	/	0.34	0.19	-	/	-	-	0.21	/	0.11	-	0.15	-	/	0.51	/	/
2013_0572	0.1	/	/	0.18	-	0.18	-	-	0.31	0.16	0.42	-	/	-	0.42	0.3	/	0.1
2015_0561	0.24	0.54	0.38	0.54	-	0.24	-	-	0.54	0.52	0.54	-	/	-	0.52	0.52	0.54	0.24
2017-0490	0.29	0.21	0.14	0.28	-	0.29	-	-	0.2	0.2	0.21	-	0.65	-	0.24	0.24	0.21	0.15
2019_0568	0.12	0.12	0.38	0.38	-	0.6	-	-	0.12	/	0.49	-	/	-	0.68	0.48	0.68	0.53

*/ no peaks, - no data

The rules are learned consecutively using training data one by one. Every time a rule is learned, the tuples that cover the rules are removed. The aim is to create a model that predicts the value of the target variable by studying simple decision-making rules derived from data characteristics. The main advantage of choosing this method is the simplicity to understand and the ability to visualize the result. On the other hand, the challenge for the implementation of the decision tree is the attributes selection referring to identifying the feature that can be considered as the root node at each level. The criterion for measuring the cleavage threshold was the Gini index [17], an indicator of the inequality of the distribution of some value of numbers. It calculates the probability of a specific feature that is classified incorrectly when selected randomly. The strategy used to select the split in each node is to find the best distribution.

2.4 Evaluation metrics

Classification predictive modeling algorithms are evaluated based on their results. In this study, we used the metric called 'accuracy' to estimate the quality of the fitted model. The accuracy of the model was determined by cross-validation, according to which the training set was randomly divided into 3 parts [18]. Each of these parts in turn acted as a test. That is, the classifier fitted three times on 3 different data sets. For each case, the accuracy of the test and training sets was calculated and averaged. The analysis of the values of these metrics made it possible to assess the accuracy, adequacy, and availability of overfitting. It means that the decision tree algorithm continues to go deeper and deeper to reduce the training-set error but results in an increased test-set error. It further reduces the accuracy of prediction in the model.

3. Results and Discussion

As can be seen from Figure 1, the most important classification features are proton density, differential proton flux in the range of 310-580 keV, and ion temperature.

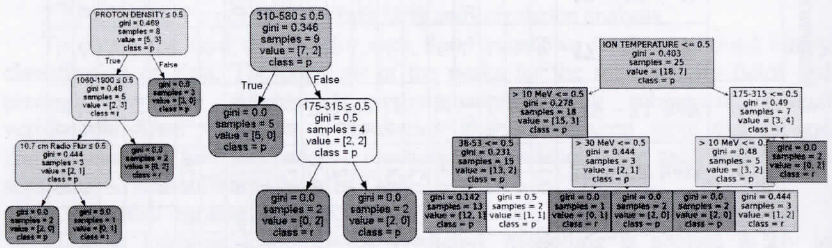

Figure 1: *The most important classification features in the forecasting of precipitation model*

Table 3 shows measures of the model accuracy for precipitation amount causing floods forecasting. The model accuracy score with criterion Gini index is higher than 0.7 in 85% of cases (17 of 20 floods). The lowest accuracy score of 0.25 was obtained for flood 2002_0463, for which no flood cause was listed in the EM-DAT database.

Table 3. *Model accuracy measures for the forecasting of precipitation*

No.	Flood events	Model accuracy score with criterion Gini index	Training-set accuracy score	Training set score	Test set score
1.	**2001_0645**	1.0000	0.8750	0.8750	1.0000
2.	**2002_0463**	0.2500	0.8571	0.8571	0.2500
3.	**2002_0480**	0.8000	1.0000	1.0000	0.8000
4.	**2002_0774**	0.8750	0.9286	0.9286	0.8750
5.	**2004_0423**	0.8750	0.9286	0.9286	0.8750
6.	**2007_0201**	0.8750	0.9286	0.9286	0.8750
7.	**2007_0247**	0.5714	0.8333	0.8333	0.5714
8.	**2007_0278**	0.8000	1.0000	1.0000	0.8000
9.	**2008_0381**	1.0000	1.0000	1.0000	1.0000
10.	**2008_0055**	0.8750	0.9286	0.9286	0.8750
11.	**2009_0497**	0.8000	1.0000	1.0000	0.8000
12.	**2012_0446**	0.8000	0.9000	0.9000	0.8000
13.	**2012_0488**	0.0000	1.0000	1.0000	0.0000
14.	**2012_0548**	0.7500	1.0000	1.0000	0.7500
15.	**2012_0549**	0.7500	1.0000	1.0000	0.7500
16.	**2012_0552**	0.6923	0.8800	0.8800	0.6923
17.	**2013_0572**	0.7143	0.9231	0.9231	0.7143
18.	**2015_0561**	1.0000	1.0000	1.0000	1.0000
19.	**2017_0490**	1.0000	1.0000	1.0000	1.0000
20.	**2019_0568**	0.7500	1.0000	1.0000	0.7500

The potential explanation of the mechanism that could explain the considered interaction in this paper was presented by several authors. According to [19] high-energy particles that come to us from the Sun capture air masses by hydrodynamics and directly affect atmospheric processes. If there is saturation with moisture at the point of contact with air masses, clouds can form and precipitation can occur, while the mechanism of precipitation formation is explained by the principle of electron valence.

According to [19] the appearance of clouds and precipitation is conditioned primarily by the electromagnetic characteristics of the solar wind, the location of the Sun from which it is emitted, and its chemical structure. According to [19] down-going atmospheric gravity waves can trigger the formation of a series of convective cells which caused heavy precipitation and floods. Consistent with these previously published results, the statistical results presented in this study show that precipitation-induced flood events tend to follow arrivals of sudden fluxes of solar-charged particles.

4. Conclusion

Even with the lack of an explanation of physical mechanisms, the establishment of an appropriate hidden dependency relationship that allows consideration of the impact of solar activity on environmental processes, such as precipitation-induced flood events, represents a contribution to this area of research. Consistently with previously published results, this study, using the Machine Learning Classification Predictive Modeling approach provided evidence that precipitation-induced flood events in the UK tend to follow the outbreak of solar wind. It is indicated that based on the detection of sudden fluxes of solar-charged particles, it is possible to expect the occurrence of precipitation leading to floods for up to several days in advance. It is found that the most important factors for flood forecasting are proton density, differential proton flux in the range of 310-580 keV, and ion temperature. Research in this paper has shown that the classification model is accurate and adequate to predict the appearance of precipitation-induced floods.

REFERENCES

1. *Milovanović B., Radovanović M.* The Connections between Solar Activity and the Circulation of Atmosphere in the 1891–2004 period // Journal of the Geographical Institute "Jovan Cvijić" SASA. 2009. № 59/1. C. 35-48. (in Serbian).

2. *Landschieidt T.* Solar wind near Earth: indicator of variations in global temperature // European Space Agency Special Publication. 2000. № 463. C. 497–500.

3. *Radovanović M., Lukić V., Todorović N.* Heliocentric Electromagnetic Long-Term Weather Forecast and its Applicable Significance // Journal of the Geographical Institute "Jovan Cvijić" SASA. 2005. № 54(1). C. 5-18.

4. *Veretenenko S., Thejll P.* Effects of energetic solar proton events on the cyclone development in the North Atlantic // Journal of Atmospheric and Solar-Terrestrial Physics. 2004. № 66(5). C. 393–405.

5. *Veretenenko S., Thejll P.*Influence of energetic Solar Proton Events on the development of cyclonic processes at extratropical latitudes // Journal of Physics: Conference Series. 2013. № 409. C. 012237.

6. *Stozhkov Yu I., Pokrevskij P.E., Okhlopkov, V.P, Zullo Zh., Martin I.M., Pellegrino Zh.K., Pinto Kh.S., Bezerra P.S., Turnelli A.* Effect of charged particle fluxes on intensity of precipitations; Russian Federation. 1996. C. 211-216

7. *Prikryl P., Rušin V., Prikryl E.A., Štastný P., Turna M., Zelenáková M.* Heavy rainfall, floods, and flash floods influenced by high-speed solar wind coupling to the magnetosphere-ionosphere-atmosphere system // Annals of Geophysics. 2021. № 39. C. 769–793.

8. *Lilensten J., Bornarel J.* Space Weather, Environment and Societies. Springer, Dordrecht. 2006.

9. *Todorovic N, Vujovic D.*_Effect of solar activity on the repetitiveness of some meteorological phenomena // Advances in Space Research. 2014. № 54. C. 2430–2440.

10. CRED Centre for Research on the Epidemiology of Disasters EM-DAT database –floods in Great Britain. (21.10.2021).

11. *NOAA Space Environment Center.* Differential Flux particles. Available at https://sohoftp.nascom.nasa.gov/sdb/goes/ace/daily/ (22.12. 2021).

12. *LISIRD, LASP Solar Irradiance Data Center.* Daily flux values. https://lasp.colorado.edu/lisird/data/noaa_radio_flux/ (22.12. 2021).

13. *Government of Canada.* Daily flux values. http://www.spaceweather.gc.ca/solarflux/sx-5-flux-en.php (22.12. 2021).

14. *Hartmann D.L.* Compositing or Superposed Epoch Analysis. ATM 552 Notes http://www.atmos.washington.edu/~dennis/552_Notes_2.pdf (20.08.2011).

15. *Cotterill, D., Stott, P., Christidis, N., Kendon, E.* Increase in the frequency of extreme daily precipitation in the United Kingdom in autumn // Weather and Climate Extremes. 2021. № 30 C. 100340.

16. *Kendon M., McCarthy M., Jevrejeva, S., Matthews A., Sparks T. Garforth J.* State of the UK Climate 2020 // International Journal of Climatology. 2021. № 41(2). C. 1–76.

17. *Raileanu L.E., Stoffel K.* Theoretical Comparison between the Gini Index and Information Gain Criteria // Annals of Mathematics and Artificial Intelligence. 2004. № 41, C. 77–93.

18. *Berrar, D.* Cross-Validation. Encyclopedia of Bioinformatics and Computational Biology. 2019. № 1 C. 542-545

19. *Stevancevic M.* Theoretic Elements of Heliocentric Electromagnetic Meteorology. Belgrade, 2006. (in Serbian).

Abstract. This paper investigates hidden dependencies between the flow of particles coming from the Sun and 20 flood events in the United Kingdom (UK). The dataset analyzed in the study contains historical data covered on the daily level for the period October 2001 – December 2019. Solar activity parameters were used as model input, while rainfall data 10 days before and during each flood event were used as model output. To determine the degree of randomness for the time series of input and output parameters the correlation analysis has been performed. Machine Learning Classification Predictive Modelling is then applied to try to establish an eventual link between input and output data. Specifically, the decision tree, as the machine learning approach is used. In addition, it is analyzed the accuracy of classification models forecast. It is found that the most important factors for flood forecasting are proton density, differential proton flux in the range of 310-580 keV, and ion temperature. Research in this paper has shown that the classification model is accurate and adequate to predict the appearance of precipitation-induced precipitation.

Keywords: solar activity, precipitation, floods, machine learning, classification, modeling