Hurricane forecasting using by parallel calculations & Mashine learning

Prof. Yaroslav Vyklyuk, DrSc, PhD.
Computer System and Technologies dep,
PHEI “Bukovinian university”
Chernivtsi, UKRAINE
vyklyuk@ukr.net

Milan M. Radovanović
Geographical Institute “Jovan Cvijić”,
Serbian Academy of Sciences and Arts
Belgrade, Serbia
m.radovanovic@gi.sanu.ac.rs

Sydor Petro
Computer System and Technologies dep,
PHEI “Bukovinian university”
Chernivtsi, UKRAINE
sydor.pete@gmail.com

***Abstract*— This research is devoted to determinate the causal relationship between the flow of particles that are coming from the Sun and emergence of the hurricanes Irma, Jose, and Katia. Five parameters i.e. characteristics of solar activity (Radio Flux 10.7, the flows of protons and electrons with maximum energy, speed of solar wind particles, and density of solar wind particles) were chosen as model input, while wind speed and air pressure of Irma, Jose, and Katia hurricanes were used as model output. Input data were sampled to six hours interval in order to adapt time interval to the observed data about hurricanes, in the period between 28 September and 21 December 2017. As a result of the preliminary analysis of 12 274 264 linear and 594 Neural Networks models using by parallel calculations, the six of them were choose as bests. The identified lags were the basis for refinement of models with the artificial neural networks. Multilayer perceptrons with back propagation have been chosen as common used artificial neural networks which are based on learning from real data. Comparison of the accuracy of both linear and artificial neural networks results confirmed the adequacy of these models. Sensitivity analysis has shown that Radio Flux 10.7 has the greatest impact on the wind speed of the hurricanes. Despite low sensitivity of pressure to change the parameters of solar wind, their strong fluctuations can cause a sharp decrease in pressure, and therefore the appearance of hurricanes.**

Keywords—Neural Network; parallel calculations; mashine learning

I. INTRODUCTION

Hurricanes endanger people's lives, destroy biodiversity, and cause enormous property damage. This requires early and reliable prediction of the emergence of the tropical cyclones. Considering that the Sun affects all biological and physical processes on the Earth, the dependence between solar activity and occurrence of Irma, Jose and Katia hurricanes was researched in this paper.

In paper [1] authors have tried, using the ANFIS model, to determine if there is a mathematical connection between the flow of high energy particles from the sun and the appearance of the hurricanes. For the period 1999-2013 (daily values from May to October), with a phase shift of 0-3 days, it was found that the models can explain at best 22%-26% of the potential connectivity. In another attempt [2], for the same time period, using better computer equipment and extending the phase shift from 0-10 days, get better results (up to 39%). The authors conclude that the results obtained cannot be ignored and that additional efforts are needed to explain the cause-and-effect relationships. In that sense, we considered that it would also be necessary to examine the causal relationship between the flow of particles from the Sun and the formation of hurricanes Irma, Katia and Jose.

In this paper an artificial neural network was established in order to learn about relationship between parameters of solar activity and data on hurricane phenomenon, measured in the period between 28 September and 21 December 2017. The modeled data fits very well with real data, especially for the wind speed of hurricanes Irma and Jose. Sensitivity analysis was conduct to determine which parameters of solar wind has the greatest impact on wind speed and pressure of Irma, Jose and Katia hurricanes. Base on the results of this research it is confirmed that it is possible to predict the appearance of hurricanes in 2-4 days ahead, after the outbreak of solar wind.

II. INPUT DATA ANALYSIS

The data on the hurricanes Irma, Jose, and Katia were downloaded from http://weather.unisys.com/hurricane/atlantic/2017/index.php. The data included maximum sustained winds in knots, and central pressure in mbar for the periods of 6 hours (0–6 hr, 6–12 hr, 12–18 hr, and 18–24 hr). The 5–minutes data on solar particle and electron flux (source: GOES-15) were downloaded from ftp://ftp.swpc.noaa.gov/pub/lists/particle/. The particles are protons (P) at > 1 Mev, > 5 Mev, > 10 Mev, > 30 Mev, and > 50 Mev. The data on electrons (E) included > 0.8 Mev, and > 2.0 Mev. The daily radio flux 10.7 data were downloaded from ftp://ftp.swpc.noaa.gov/pub/indices/old_indices/2017Q3_DSD.txt. The data on proton speed (km/s) and proton density (protons per cubic centimetre) were downloaded http://umtof.umd.edu/pm/crn/. The source of the data was SOHO CELIAS Proton Monitor.

Each record represents an averaged metric for the specified sampling interval. The dynamics of the mentioned above characteristics is shown in Figures 1a and 1b.

III. PRELIMINARY PROCESSING OF INPUT DATA

The range of input parameters is greater than the output ones. It is allows us to take into account lag dependencies without reducing the number of the time series. It should be noted that the sampling of the measurement of values in all cases except

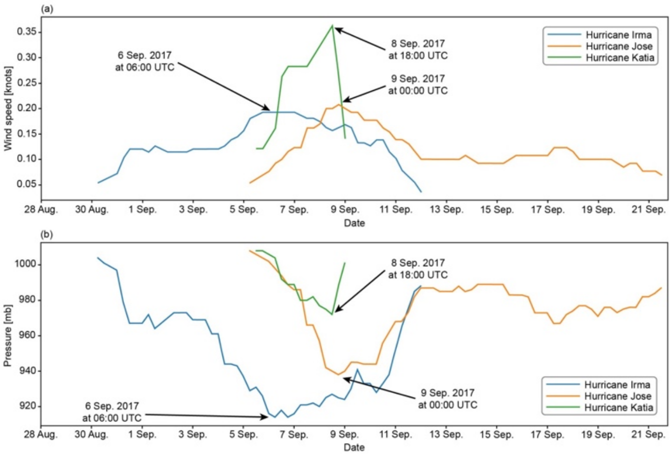

Fig. 1. Wind speed (a) and pressure (b) for hurricanes Irma, Jose and Katia. Black arrows represent dates of maximum wind speed and air pressure respectively.

Radio Flux 10.7 is greater than the researched output values. For the further research the sampling of all input data was reduced to six hours by averaging. The averaging was made taking into account that the value at a given moment of time is averaged over the entire previous period between measurements, not including measurements at a given time.

In the case of the Radio Flux time series, whose sampling is greater than the output fields, the data was interpolated with a sampling of 6 hours. Cubic spline interpolation using Hermite polynomials (PCHIP) was used [3].

IV. CORRELATION ANALYSIS

In order to find the relationship between input factors it was carried out a correlation analysis. It shown a strong correlation between time series of protons flows. The situation is the same for the electron flows. So the number of input factors can be significantly reduced. In order to select the most appropriate factor, the time series was normalized and depicted on a single graph (Figure 2a, 2b, and 2c).

As can be seen from the Figure 2a, all normalized factors that describe the flow of protons are characterized by the same dynamics. Namely, all 6 time series have two distinct peaks. The first peak corresponds to the date of 7 September 2017, the second to 11 September 2017. Namely, the first came later after the wind speed extreme hurricane Irma and somewhat ahead of the hurricanes Jose and Katia. The second peak appears after all three hurricanes, so it is unlikely that the flow of protons affects the appearance of hurricanes.

As can be seen from the Figure 2b, the behavior of electron fluxes of different energies is quite similar. There are pronounced oscillations on the graphs that are not visually observed in the dynamics of the studied characteristics of hurricanes. The behavior of the SW (Figure 2c) also significantly differs from Figure 1. That is why there is low probability of influence of these factors on wind power and hurricane air pressure.

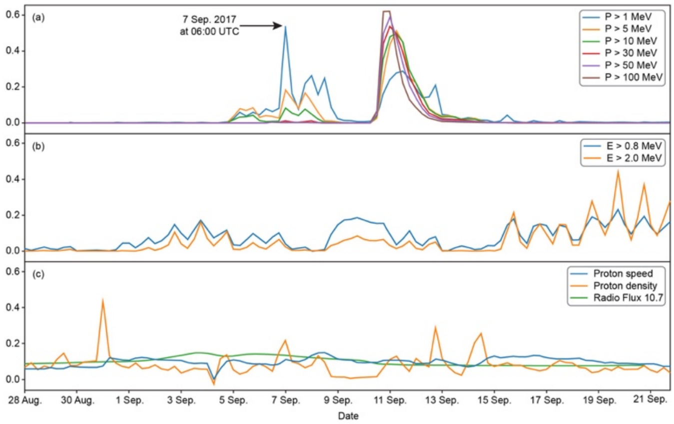

Fig. 2. Normalized input parameters of proton flows (a), electron flows (b), speed, density, and Radio Flux (c).

In order to confirm or refute these conclusions a lag correlation analysis was conducted, which allowed to find a correlation between the separate time series of input factors displaced for a certain number of rows vertically down (lag) and output factors [4].

The highest correlation coefficient R is observed for the factor Radio Flux 10.7 for the Irma hurricane: Rwind speed = 0.86 (lag = 6), Rpr = –0.91 – pressure (lag = 9). Then Katia (lag = 17) Rwind speed = 0.84, and Rpr = –0.91. The hurricane Jose has the smallest correlation coefficient (lag = 18) Rwind speed=0.72 and Rpr = –0.47. Negative values of correlation coefficients for pressures time series prove an inverse relationship between the input factor and the output one. This confirms the preliminary conclusions about the relationship between this factor and hurricane parameters.

Other factors also have rather high correlation coefficients, including the electron flux and the characteristics of the SW. This refutes the previous hypothesis of the absence of their impact on hurricanes. The flow of protons has a high correlation coefficient only for the Katia hurricane. This may be explained by the small number of observations for this hurricane.

In addition, the distribution of lags with maximum (minimum) correlation coefficients is significant for different input variables. Thus, high correlation coefficients and the unresolved issue of lags causes the further research to find functional relationships between these input and output parameters. Therefore, Radio Flux 10.7, the flows of protons and electrons with maximum energy (P > 100 and E > 2.0) and speed and density of solar wind particles were selected as test parameters for the further research.

V. PARALLEL CALCULATIONS FOR FINDING OPTIMAL MODELS

For easy formalization of the models we combine the output time series into the target vector and the input parameters into the vector of parameters:

$$T = (T_1, T_2, T_3, T_4, T_5, T_6), \quad (1)$$

$$X = (X_1, X_2, X_3, X_4, X_5), \quad (2)$$

where T_i is time series of the wind speed and the pressure of Irma, Jose, and Katia hurricanes respectively; X_i is time series of P > 100, E > 2.0, speed of solar wind particles, density of solar wind particles, and Radio Flux 10.7 respectively.

The task is to find for each T_i the most accurate and adequate functional dependence of the type:

$$T_i = F_i(X, L_i, \Omega_i), \quad (3)$$

where $L_i = \{l_{ij}\}_{j=1,5}$ is the vector of optimal lags and Ω_i is parameter of the linear or the artificial neural network model.

Determination coefficient (R^2) and mean square error (MSE) were considered as criteria of optimization. As further calculations have shown, in most cases the model with the maximum determination coefficient and the minimum mean square error coincided. In the case of differences in these models, models with the maximum value of the determination coefficient were more adequate. Therefore, the further research uses the determination coefficient as a criterion of adequacy.

Cross-validation for k-blocks (k-fold cross-validation) was used to verify the accuracy of the models [5]. The training sample was divided into *k* blocks of the same size. Each block alternately was as a test sample, and the other *k-1* blocks were a training sample. The result of adequacy was determined by calculating the value of the determination coefficient between the components of the target vector T_i and the model data on the values of the test samples $F_i^{cv}(X, L_i, \Omega_i)$. The size of the test sample was chosen as 10% of the total size of the training sample, i.e. k = 10. In general, the optimization problem has the form:

$$R_i^2\left(T_i, F_i^{cv}(X, L_i, \Omega_i)\right) \xrightarrow{yields} max \quad (4)$$

Solution variables: $L_i \in Tasks, \Omega_i$

Limitations: $l_{ij} < 22;$

$$\max_{j=1,5}\{l_{ij}\} < (lag - 2)_{Lag=\overline{0-22}},$$

where Ω_i is the parameter of the model, which is determined by fitting the initial model data to the target vector, the fitting method depends on the type of model (linear, neural network, etc.). The optimization was done by completely scanning all possible combinations of the lag vector L_i for each component of $X_i = \{x_{ij}\}_{j=1,5}$ from 0 to 22. The magnitude of the maximum lag was chosen from the preliminary analysis where the maximum lag was 20. Therefore, it was decided to check two lags more. In this case, the set of lag combinations is defined as the Cartesian product of the test lag vectors for each input parameter and is 23^5 = 6,436,343:

$$Tasks(22) = \prod_{j=1,5} L(22), \quad (5)$$

where $L(22) = \{0,1,\dots,22\}$.

For cross-validation it is necessary to optimize 10 models plus one additional for a full set of values of the training sample. Such optimization should be performed for each element of the target vector, which is six. With this in mind, the total number of models that need to be optimized will be: $23^5 \cdot 11 \cdot 6$ = 424,798,638. Such a huge number of tasks require an optimal choice both for the type of model (see equation (3)) and optimization algorithms.

In order to reduce the number of tested lags, an algorithm for finding an optimal model was proposed:

1. The first maximum number of lags is determined *lag = 0*.
2. A set of tasks is formed based on the equation (6):
3. For the first run *Tasks (lag)*.
4. For the next runs in order to avoid repetitions of tasks the difference of sets needs to be calculated *Tasks (lag) '= Tasks (lag) \ Tasks (lag-1)*.
5. The optimal model is found according to equation (5).
6. If the maximum lag value for any component of the optimal model does not exceed *lag-2*, it is assumed that the optimal value is found and the algorithm is completed.
7. If *lag = 22*, the algorithm is completed and is considered to have no optimal value.
8. Increase *lag + = 1* and move to step 1.

The presence of the set of several independent tasks which use the same memory area causes using parallel calculations by forming a pool for multiprocessing tasks, according to equation (4). After calculating all tasks, a function with a maximum determination coefficient was found and the conditions 4 and 5 of the algorithm were checked [6].

Linear models were selected as test cases. This allowed reduction of the calculation time and determination of the optimal lags for each of the input parameters for all goal vectors.

VI. Refinement of Models Using Artificial Neural Networks

As a result of the preliminary analysis, six linear models are obtained using the equation (3). That is why the parameters of the linear models Ω_i т and the lags of the input parameters L_i are known. The identified lags were the basis for refinement of models with the artificial neural networks. The maximum size of training samples is 66, and a minimum is 15 records. Each neuron network must have five inputs. Such a small size of the training sample puts some restrictions on the size of the neural networks and the possibility of their adequate training. Multilayer perceptrons (MLPs) with back propagation were chosen [7]. The method of training was as a quasi-Newtonian method of optimization. Logistic function was selected as an activation function.

As test calculations have shown, the best results were observed for single-layered neural networks with the number of neurons in a hidden layer, which equals to seven. Decreasing the number of neurons in the hidden layer reduced the network's ability to learn. Increasing the number of neurons in the hidden layer, on the contrary, led to retraining. Namely, on the one hand the results of learning on training set have improved significantly, but on the other side significant fluctuations in the results of the cross-validation test appeared. However, even in the best of cases, there were from one to two abnormal fluctuations in the results of models, which disappeared during repeated training in one place of the time series and appeared in another. To remove these fluctuations, Delphi expert valuation method was used [8]. Its essence was:

1. Several neural networks were created and studied for each model according to equation (3). In our case, their optimal number was nine. Their increase did not improve the result.
2. Predictive values were calculated on the test sets of data using the cross-validation method for each of the networks. The result was a matrix of type:

$$Res_i = \begin{bmatrix} f_{i1}^1 & \cdots & f_{im}^1 \\ \vdots & \ddots & \vdots \\ f_{i1}^9 & \cdots & f_{im}^9 \end{bmatrix}, \quad (6)$$

where m is the size of the training sample for a particular vector of the goals, the upper index is the serial number of the neural network.

3. Each column was sorted and then 10% of records with minimum and maximum values were removed from the records.
4. For the remaining values for each of the columns the median was determined, which was considered to be the result.

As a result of this phase, unlike linear models, for each of the target vectors a list of nine trained neural networks was received: $T_i = \{F_i^n(X, L_i, \Omega_{i,n}^{ANN})\}_{\overline{n=1-9}}$. The total number of neural networks was: $6 \cdot 9 \cdot 11 = 594$, considering that for each cross-validation, 10 neural networks + 1 network were constructed and studied on the complete training sample (necessary for further sensitivity analysis). To reduce computer time, parallel calculations were also used (the same as at the cross-validation level).

VII. PARALLEL CALCULATIONS RESULTS OF ARTIFICIAL NEURAL NETWORKS AND LINEAR MODELS

The best way to describe how modeled data fit to the real data is to plot them on the single graph for each hurricane that has been subject of this study. The results of calculations for linear models and artificial neural networks are presented in the Figure 3

As can be seen from the Figure 3, in the majority of cases, linear models show the best prediction result for all six target vectors under study. The best results for neural networks are obtained for the Irma and Jose hurricanes (i.e. wind speed). As can be seen from the Figures 3d–3f neural networks show significantly worse results than linear models. In the case of the Katia hurricane, the lag of the optimal model for the RadioFlux field, which is four and zero, is also strange. This is completely inconsistent with the previous analysis. This can be explained by the small dataset size (15 records) and, accordingly, the inability of adequate training for both linear and neural networks. Regarding the Irma and Jose hurricanes, the obtained lags are in good agreement with the previous analysis for the RadioFlux field, which is the most influential (as it has been shown above).

A quantitative comparison of the accuracy of the results is presented in the Table 1.

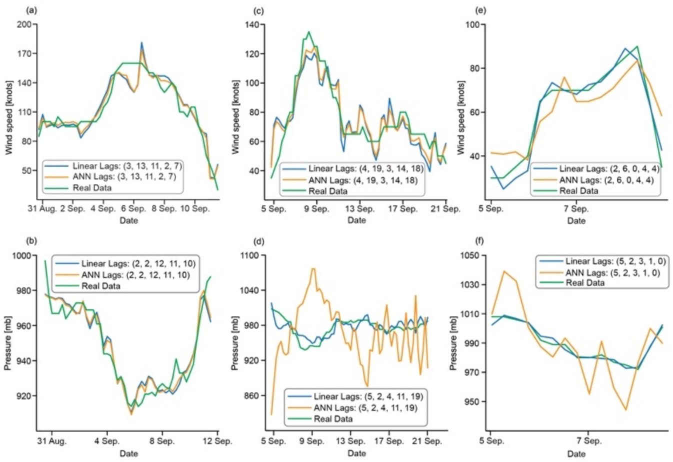

Fig. 3. Results of hurricane forecasting with linear models and artificial neural networks

TABLE I. LAGS AND CORRELATION COEFFICIENTS OF THE OBTAINED MODELS

Hurricane	Parameter	Model: Equation	Model: Type	Numbers of tests models	Lags	R^2 Full dataset	R^2 Cross validation
Irma	Wind speed	$F_1(X, L_1, \Omega_1^{Lin})$	Linear	1,048,576	L_1 =(3, 13, 11, 2, 7)	0.89	0.85
		$\{F_1(X, L_1, \Omega_1^{ANN})\}$	ANN	99		0.89	0.75
	Pressure	$F_2(X, L_2, \Omega_2^{Lin})$	Linear	759,375	L_2 = (2, 2, 12, 11, 10)	0.90	0.88
		$\{F_2(X, L_2, \Omega_2^{ANN})\}$	ANN	99		0.90	0.87
Jose	Wind speed	$F_3(X, L_4, \Omega_4^{Lin})$	Linear	5,153,632	L_3 = (4, 19, 3, 14, 18)	0.86	0.77
		$\{F_3(X, L_4, \Omega_4^{ANN})\}$	ANN	99		0.86	0.74
	Pressure	$F_4(X, L_4, \Omega_4^{Lin})$	Linear	5,153,632	L_4 = (5, 2, 4, 11, 19)	0.69	0.56
		$\{F_4(X, L_4, \Omega_4^{ANN})\}$	ANN	99		0.61	0.62
Katia	Wind speed	$F_5(X, L_5, \Omega_5^{Lin})$	Linear	100,000	L_5 = (2, 6, 0, 4, 4)	0.98	0.96
		$\{F_5(X, L_5, \Omega_5^{ANN})\}$	ANN	99		0.72	0.34
	Pressure	$F_6(X, L_6, \Omega_6^{Lin})$	Linear	59,049	L_6 = (5, 2, 3, 1, 0)	0.98	0.96
		$\{F_6(X, L_6, \Omega_6^{ANN})\}$	ANN	99		0.65	0.53
Total			Linear	12,274,264			
			ANN	594			

The Table 1 shows that the highest correlation coefficients are obtained for target vectors such as the wind speed of the Irma hurricane, the pressure of the Irma hurricane, and the wind speed of the Jose hurricane. The determination coefficients for linear models and neural networks coincide. Cross-validation results are slightly lower, but they also have high values. This also confirms the adequacy of these models. The table also demonstrates that pressure of the Jose hurricane has low values of correlation coefficients for both linear models and for neural networks. Therefore, the accuracy of this model is low. Regarding the Katia hurricane, it should be noted that similarly to the graphs, the results are accurate for linear models and low for neural networks, which may be caused by the small amount of the training sample.

During the calculations, the following optimal linear models were obtained:

$$F_1(X, L_1, \Omega_1^{Lin}) = -16.44 - 1.09 \cdot x(3)_1 + 2.88 \cdot 10^{-04} \cdot x(13)_2 - 0.05 \cdot x(11)_3 + 0.85 \cdot x(2)_4 + 1.40 \cdot x(7)_5,$$

$F_2(X, L_2, \Omega_2^{Lin}) = 1067.52 + 0.55 \cdot x(2)_1 - 5.42 \cdot 10^{-04} \cdot x(2)_2 + 0.02 \cdot x(12)_3 + 0.63 \cdot x(11)_4 - 1.17 \cdot x(10)_5$,

$F_3(X, L_3, \Omega_3^{Lin}) = -80.15 - 0.71 \cdot x(4)_1 + 4.93 \cdot 10^{-04} \cdot x(19)_2 + 0.12 \cdot x(3)_3 + 1.62 \cdot x(14)_4 + 0.84 \cdot x(18)_5$,

$F_4(X, L_2, \Omega_2^{Lin}) = 1073.42 + 0.54 \cdot x(5)_1 - 2.83 \cdot 10^{-04} \cdot x(2)_2 - 0.08 \cdot x(4)_3 - 1.27 \cdot x(11)_4 - 0.52 \cdot x(19)_5$,

$F_5(X, L_5, \Omega_5^{Lin}) = -413.61 - 94.62 \cdot x(2)_1 - 8.08 \cdot 10^{-04} \cdot x(6)_2 + 0.17 \cdot x(0)_3 - 1.88 \cdot x(4)_4 + 3.14 \cdot x(4)_5$,

$F_6(X, L_6, \Omega_6^{Lin}) = 783.42 - 26.24 \cdot x(5)_1 + 1.42 \cdot 10^{-04} \cdot x(2)_2 + 0.12 \cdot x(3)_3 - 2.30 \cdot x(1)_4 + 1.19 \cdot x(0)_5$.

where the value of the lag is specified in the brackets of the input parameters.

The result of the training neural networks is represented by 54 neural networks whose parameters change during the training, so it is not expedient to bring such a number of dynamic matrices of neurons' weight factors.

The total number of tested linear models on account of the proposed algorithm using decreased from 424,798,638 to 12,274,264 · 11 = 135,016,904, i.e. the total number of models is reduced by 3 times and consists 32% of the previous indicator. Considering that the calculation time was 4.5 hours on the Mac Book Pro (2015), it saved approximately 14 hours of computer time.

The using of neural networks within such an algorithm takes several orders more time and requires, accordingly, the involvement of a computer cluster.

VIII. Sensitivity Analyses

In order to verify the adequacy of the models an analysis of the model's sensitivity to the change of factors for all models in the Table 5 was performed [9]. The analysis was as follows. For each tuple r of the input parameters vector $X^r = \{x_j^r\}_{\substack{j=\overline{1-5},\\ r=\overline{1-N}}}$, which consists of N records, the input parameters value was incremented by 10% and the change of the corresponding model $F_{i=\overline{1-6}}$ or the set of models was calculated by the Delphi method (in the case of neural networks). Then all the obtained values were averaged. The resulting value means the average change in wind speed or pressure of the particular hurricane with an increase of the input parameter by 10%.

In order to implement this a diagonal matrix of variation factors was created with dimension which equals to the number of input parameters, in our case, five:

$$V = \begin{bmatrix} 0.1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0.1 \end{bmatrix}_{5\times 5}. \tag{7}$$

Each tuple of the input parameters vector is duplicated vertically in the amount which equals to the length of the tuple (that is, the number of input parameters):

$$A^r = \begin{bmatrix} x_1^r & \cdots & x_5^r \\ \vdots & \ddots & \vdots \\ x_1^r & \cdots & x_5^r \end{bmatrix}. \tag{8}$$

The matrix of test values is calculated as an elemental product of matrices:

$$T^r = (V+1) \cdot A^r = \begin{bmatrix} 1.1 \cdot x_1^r & \cdots & 1.0 \cdot x_5^r \\ \vdots & \ddots & \vdots \\ 1.0 \cdot x_1^r & \cdots & 1.1 \cdot x_5^r \end{bmatrix} \tag{9}$$

The vector of values is calculated:

$$S_i^r = F_i(T^r, L_i, \Omega_i^{Lin(ANN)}) = \begin{bmatrix} f_{i,x_1}^r \\ \vdots \\ f_{i,x_5}^r \end{bmatrix}. \tag{10}$$

An array of obtained changes of functions F_i is formed by estimating the values S_i^r for all tuples of the vector X:

$$S_i = \begin{bmatrix} (S_i^1)^T \\ \vdots \\ (S_i^N)^T \end{bmatrix}. \tag{11}$$

Then the vector of predicted by the model values is calculated and duplicated horizontally by the amount of input fields:

$$M_i = \{m_i^r\}_{r=\overline{1-N}} = F_i(X, L_i, \Omega_i^{Lin(ANN)}), \tag{12}$$

$$Mx_i = \begin{bmatrix} m_i^1 & \cdots & m_i^1 \\ \vdots & \ddots & \vdots \\ m_i^N & \cdots & m_i^N \end{bmatrix}_{N\times 5}. \tag{13}$$

The last step is the building of matrix of relative changes by calculating the elemental difference and dividing the matrices Mx_i and S_i. Then averaging by the columns is made:

$$D = (S_i - Mx_i)/Mx_i, \tag{14}$$

$$Sens = \overline{D}_{col}. \tag{15}$$

The results of calculations are given in the Table 2.

TABLE II. SENSITIVITY ANALYSIS OF THE OBTAINED MODELS

Hurricane	Parameter	Model	P > 100	E > 2.0	Speed	Density	Radio Flux 10.7
Irma	Wind speed	Linear	–0.63%	0.10%	–2.51%	0.23%	14.38%
		ANN	–0.65%	0.13%	–2.64%	0.18%	13.05%
	Pressure	Linear	0.02%	–0.04%	0.09%	0.02%	–1.36%
		ANN	0.02%	–0.04%	0.09%	0.02%	–1.36%
Jose	Wind speed	Linear	–0.26%	0.50%	9.27%	0.64%	11.27%
		ANN	–0.26%	0.50%	9.27%	0.64%	11.27%
	Pressure	Linear	0.01%	–0.04%	–0.42%	–0.04%	–0.53%
		ANN	0.00%	0.59%	3.63%	0.43%	5.24%
Katia	Wind speed	Linear	–1.07%	–1.19%	17.69%	–1.17%	74.57%
		ANN	0.00%	–1.30%	8.80%	–0.64%	3.33%
	Pressure	Linear	–0.02%	0.05%	0.66%	–0.07%	1.46%
		ANN	0.00%	–0.14%	2.65%	0.50%	6.98%

As can be seen from the Table 2, the factor that has the greatest impact on the wind speed of the hurricanes is Radio Flux 10.7. Its increase by 10% leads to an increase in the wind speed for the Irma hurricane on average by 13%–14% in 42 hours (lag 7) and 11% in 4.5 days (lag 18) for Jose. As the table shows, indicators of linear models and neural networks are sufficiently close for all factors and these hurricanes, which confirm the adequacy of the models. The second important indicator is the speed of the SW. It's increasing by 10% raises the hurricane Jose's speed by 9% after 18 hours (lag 3) and decreases the hurricane Irma's speed by 2.5% after 3 days. Other factors do not affect these two hurricanes.

For the Katia hurricane Radio Flux 10.7 is 74% for the linear models and only 3% for neural networks. A strong difference in the sensitivity of neural network and linear models also calls into question their adequacy. This may be caused by a small amount of data, which prevented the construction of an adequate model.

As known, the root cause of the wind is the pressure drop, so it is interesting to analyze the influence of the parameters of the SW on air pressure. If we analyze the sensitivity of the pressure for Irma and Jose hurricanes, we can see that they are less sensitive to changes in SW. In particular, changing the Radio Flux 10.7 by 10% causes a pressure drop of 1.3% after 2.5 days for the Irma hurricane and practically does not affect the pressure of the hurricane Jose. However, as can be seen from the Figure 2, the indicated parameter has increased from 28 August to 4 September 2017 from 82.4 to 140, that is, by 70%. According to the Table 7, the change in only one of these factors had to cause a pressure change in the hurricane zone at $0.7 / 0.1 \cdot (-1.3\%) = -9.5\%$, that is, from 1004 mb to 908 mb. The actual recorded pressure was 914 mb (forecast error is 0.6%). For the hurricane Jose, the calculated change is 971 mb, the recorded is 938 mb (forecast error is 3.5%). Thus, despite the low sensitivity of pressure to change the parameters of SW, strong fluctuations of the input parameters can cause a sharp decrease in pressure, and hence the emergence of hurricanes.

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

IX. CONCLUSIONS

Considering the porential prognostic models, one should certainly bear in mind that for solar flares from active regions located at the East of the helilongitude, the time delay (between emission and the ground level enhancement onset) can be from several hours up to days. Almost all diffusion models involving solar particle transport in the interplanetary medium show that the maximum time delay is proportional to the square of the distance traveled [10].

The efficiency of the penetration depends on the degree to which the interplanetary magnetic field provides input of the particle flux to the region with the given angle and/or in what percent relation the particles of the given direction are present in the flux with a high angular isotropy.

Research in this paper has shown that applied model is accurate and adequate to predict the appearance of hurricanes 2–4 days ahead, after the outbreak of SW. High correlation coefficients sustain the previous conclusion. About 90% of variations of the Irma hurricane can be explained by the model. Jose is the hurricane in the Pacific Ocean, which has larger scale, and therefore the processes of the influence of external factors are more inertial, which explains a bigger lag in the calculations. The sensitivity analysis revealed that Radio Flux 10.7 has the greatest impact on wind speed of the hurricanes, except in the case of the Katia hurricane. In the general picture of the change in pressure and wind speed over a longer period, there are other factors that were not taken into account in the model. Therefore, the model for Jose was less accurate, but quite adequate. As already had been noted in the section 8, the Katia hurricane was the least lengthy and the data were not enough to test the hypothesis in this case.

The coupling of the stratosphere with surface climate is one good candidate to better understand the signals of the future climate changes [11]. Vertical wind shear was shown to be a much more fundamental component for major hurricane development and maintenance [12].

REFERENCES

[1] Vyklyuk, Y., Radovanović, M., Milovanović, B., Leko, T., Milenković, M., Milošević, Z., et al. (2017). Hurricane genesis modelling based on the relationship between solar activity and hurricanes. Natural Hazards, 85(2), 1043–1062.

[2] Vyklyuk, Y., Radovanović, M. M., Stanojević, G. B., Milovanović, B., Leko, T., Milenković, M., et al. (2017b). Hurricane genesis modelling based on the relationship between solar activity and hurricanes II. Atmospheric and Solar-Terrestrial Physics. Advance online publication.

[3] Fritsch, F. N., & Carlson, R. E. (1980). Monotone piecewise cubic interpolation. SIAM Journal on Numerical Analysis, 17(2), 238–246.

[4] Olden, J. D., & Neff, B. D. (2001). Cross-correlation bias in lag analysis of aquatic time series. Marine Biology, 138(5), 1063–1070.

[5] Trippa, L., Waldron, L., Huttenhower, C., & Parmigiani, G. (2015). Bayesian nonparametric cross-study validation of prediction methods. The Annals of Applied Statistics, 9(1), 402–428.

[6] Palach, J. (2014). Parallel Programming with Python. Birmingham, UK: Packt Publishing Ltd.

[7] Géron, A. (2017). Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems(1st ed.). Sebastopol, CA: O'Reilly Media, Inc.

[8] Chang, P. T., Huang, L. C., & Lin, H. J. (2000). The fuzzy Delphi method via fuzzy statistics and membership function fitting and an application to the human resources. Fuzzy Sets and Systems, 112(3), 511–520.

[9] Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., & Wagener, T. (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow. Environmental Modelling & Software, 79, 214–232.

[10] Augusto, C. R. A., Kopenkin, V., Navia, C. E., Felicio, A. C. S., Freire, F., Pinto, A. C. S., et al. (2013). Was the GLE on May 17, 2012 linked with the M5. 1-class flare the first in the 24th solar cycle?

[11] Kidston, J., Scaife, A. A., Hardiman, S. C., Mitchell, D. M., Butchart, N., Baldwin, M. P., Gray, L. J. (2015). Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. Nature Geoscience, 8, 433–440.

[12] Klotzbach, P. J. (2006). Trends in global tropical cyclone activity over the past twenty years (1986–2005), Geophysical Research Letters, 33(10), L10805.