A MACHINE LEARNING APPROACH FOR FLOODPLAIN MANAGEMENT AND RESTORATION ON THE LOWER DANUBE, A CASE STUDY BETWEEN BOIANU – CĂLĂRAȘI, ROMANIA

Mihnea Cristian Popa[1,2]
Mila Chilikova-Lubomirova[3]
Daniel Constantin Diaconu[1,4], *
[1] Centre for Integrated Analysis and Territorial Management,
  University of Bucharest, **Romania**;
[2] Simion Mehedinți "Nature and Sustainable Development"
  Doctoral School, University of Bucharest, **Romania**;
[3] Institute of Mechanics – Bulgarian Academy of Sciences, **Bulgaria**;
[4] Department of Meteorology and Hydrology, Faculty of Geography,
  University of Bucharest, **Romania**.

ABSTRACT

Over time, the Lower Danube floodplain has undergone many regulation and damming works or was subject to other types of human activities which gradually changed its dynamics and ultimately contributed to the altering of the existing ecosystems. In Romania, along the Lower Danube, large floodplain areas have been lost or transformed through extensive damming especially in the second of part the 20th century to obtain new agricultural lands which over the years have decreased in fertility. Since the first damming works, around 81.4% of the Lower Danube floodplain in Romania has been dammed, but since 1990 only 5% has been restored. The aim of the study is to integrate Machine Learning with GIS tools and modelling techniques for floodplain restoration to create permanent or semi-permanent wetlands which could store and be used to control high flows during floods in the lower sector of the Danube, a case study between the former Boianu pond and the Călărași municipality. The study area has a total surface of 23.000 hectares and was heavily affected during the 2006 and 2010 floods. An optimal and organized management of the floodplain would lead to the improvement of the ecosystems services and would reduce the impact of floods downstream.

Keywords: Danube, floodplain management, river restoration, machine learning, GIS.

INTRODUCTION

During the last century, the Danube floodplain has undergone many transformations, especially through damming works, so that a large extent of the existing ecosystem has been altered. The changes made on the Danube floodplain through its damming and the catastrophic floods over the years have affected the local communities both from an economic, cultural and infrastructure point of view, represent a topic addressed in previous studies [1], [2], [3], [4], [5]. The restoration of the lost Danube wetlands has also been the objective of certain organizations that have started projects on this matter, for example, the World Wildlife Fund (WWF). The end of the 19th century represents the beginning of the first damming works in Romania, the first work being carried out in the Danube Delta in 1895 at Mahmudia. Studies carried out after the damming works noted

that the new agricultural lands were fertile. By the end of the 1980s, around 84% of the entire Romanian Danube floodplain was dammed, with a maximum area reaching 430.000 hectares. Over time, the fertility of those new agricultural lands gradually decreased and the existing ecosystems were completely altered [1]. With the construction of hydropower dams and dikes, the lateral connectivity of the river with its floodplain has been minimized thus causing a loss in biodiversity. The lateral connectivity of a river with its floodplain represents an important interaction, following which an exchange of organisms, matter, sediments and other beneficial nutrients for the ecosystems is being done. All alterations brought to a riverbed or its floodplain, lead to a change of connectivity thus having multiple effects of the ecosystem.

The study area is represented by the Boianu-Sticleanu-Călărași enbankment, with an area of 23.000 hectares which overlaps 5 territorial administrative units between the former Boianu pond and the Călărași municipality (Figure 1). The first damming works in the Călărași county were carried out in the 1900s. The works for the Boianu-Sticleanu-Călărași embankment were carried out between 1962 and 1963, a period in which most of the damming works in Romania were performed. With the fall of the communist regime in Romania in 1989, about 5% of the Danube floodplain was restored. These restoration works were carried out both through national and European programs and through NGOs (Rusenski zlom and Persina, Bulgaria) [6] .The 5% achieved so far are not enough to restore the ecosystem of the floodplain.

The hydrologic regime of the Danube is characterized by periods of with floods in the spring months, especially in April and May. The floods are a result of the abundant precipitations that cause flows of the Danube tributaries to be high but also as a result of the melting of snow [7], [8].
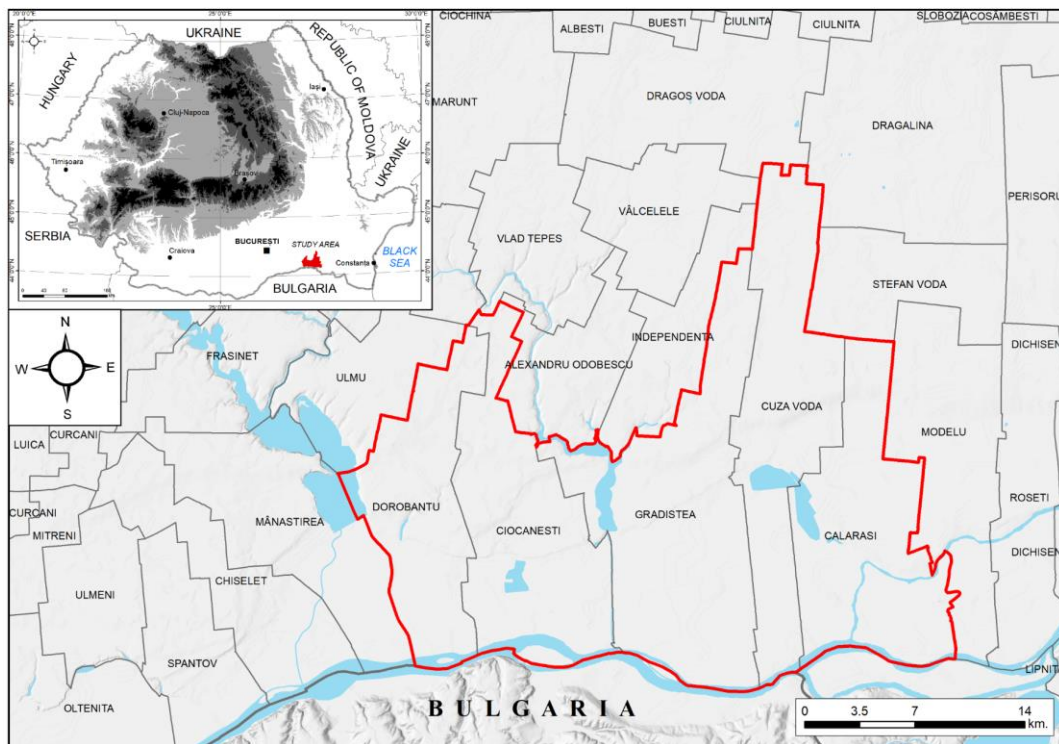

Figure 6. General location map

All these changes as a result of various human activities lead to an alteration or permanent loss of certain special in the local biodiversity [9], [10]. The objective of this study is to identify solutions for the management of the lower Danube floodplain in order to mitigate the effects of floods and restore its floodplain.

METHODOLOGY

The methodology proposed by this study uses the Multilayer Perceptron (MLP) artificial neural networks and the Frequency Ratio (FR) bivariate statistical model and their hybrid ensemble. The analysis carried out in this study uses open-source data, relevant for researchers for developing / testing methodologies.

The first and one of the most important steps in the analysis, vital for computing the models, consisted in the inventory of the potential locations for creating permanent or semi-permanent wetlands. The inventory of the locations was done using historical maps of the area from which the former wetlands that existed until the damming works of the Danube were identified. The earliest map of the area dates from 1864 and the most recent being the topographic map of Romania made between 1987 – 1993. Other maps identified date from 1910 (the 3rd Military Mapping Survey of Austria-Hungary) and 1976 (the Soviet military topographic map sets). Other sources used include the CORINE Land Cover data set from 2018, a Digital Elevation Model (DEM) with a global resolution of 15 meters, the hydrological soil groups (HSG) extracted from the pedological map of Romania. Sentinel-II satellite imagery with a spatial resolution of 10 meters from August 2018 were used in the process of inventorying the potential areas for renaturation. The slope and slope-aspect were derived from the DEM. Based on the inventory of locations, a database was created containing 130 locations encoded with the values of 1 (Potential Area) and 0 (Low Potential Area). As there isn't a perfect split ratio, the training and testing data were split into a 70-30 % ratio after multiple tests. Both sets of data contain the values of the factors which overlap the locations. These values were extracted using the *Extract Multi Values to Points* tool in ArcGIS.

The selection of the factors used to perform the analysis on the potential areas for the renaturation of the wetlands represents the next step in preparing the data for computing the models. The study proposes the use of 5 factors and are as follows: Elevation, Slope, Slope-Aspect, Land-use and the Hydrological Soil Groups. The factors were exported in a raster format with a resolution of 15 meters and were reclassified into 3 classes, determined according to the elevation, slope, slope-aspect, soil permeability and land-use.

The Multilayer Perceptron represents an artificial neural network used in pattern recognition and function approximation and is made up of three components, the input layers, the hidden layers and the output [1], [2], [3] (Figure 2). An MLP recognizes previously unseen data and can solve the problems between the presence or absence of a phenomenon. The MLP uses the backpropagation algorithm which is a supervised learning technique [4], [5], [6]. The input layers also known as the neurons are connected to the hidden layers through a series of neural connections. The neural connections hold the weights of the hidden layers. The hidden layers are connected to the output layers through neural connections which hold the weights of the output layers. The weights of the neural connections are random until they intersect and are multiplied by the value of that intersection. [11], [16], [17]. The MLP model was trained using 500 training epochs and 30 validation thresholds. The validation thresholds allow the algorithm to check for a decrease in error in the neural network [11].
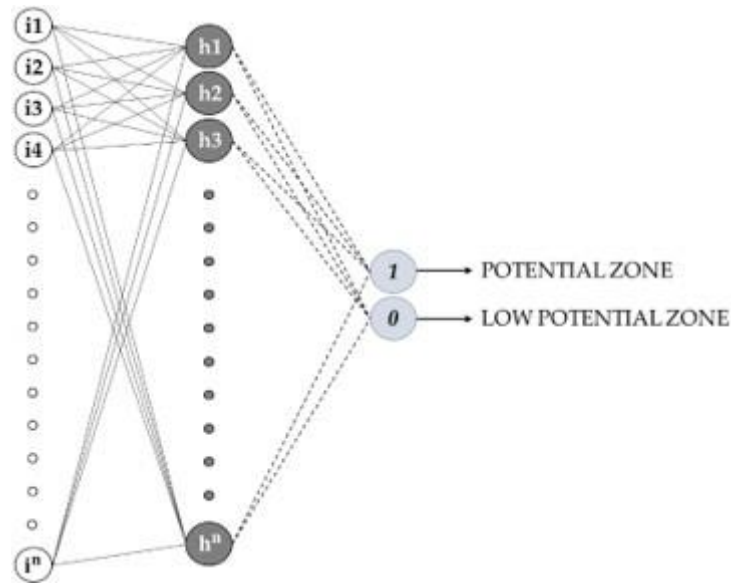
Figure 7. Multilayer Perceptron neural network diagram

The Frequency Ratio (FR) represents a bivariate statistical model used to determine the ratio of an area in which the probability ratio of an occurrence to a non-occurrence and the occurrence of a phenomenon is present in a dataset [18]. The FR is determined based on the relation between the training / testing locations and the selected geographical factors. The high prediction ratio (PR) values show that the factor has a high influence on a certain zone for it to be suitable to be renatured. The PR was determined using equation (1), where *SA* is the maximum and the minimum spatial association between the training / testing locations and the selected factors [17], [19].

$$PR = \frac{(SA_{max} - Sa_{min})}{(SA_{max} - Sa_{min})\,min} \tag{1}$$

After determining each PR value for the selected geographical factors, each raster was reclassified using the relative frequency values (RF), equation 2:

$$RF = \frac{R_+}{R_{tot}} \tag{2}$$

where R_+ represents the positive ratio and R_{tot} is the sum of each R_+.
After each raster was reclassified, the FR model was computed using equation 3, where *n* represents the number of geographical factors and W_{ij} is the weight of class *i* pf the parameter *j* [20].

$$FR = \sum_{j=1}^{n} W_{ij} \tag{3}$$

The MLP-FR hybrid ensemble was computed in ArcGIS using the *Raster calculator* tool through the integration of the R_+ values and the weights resulted from the MLP neural network (Figure 3) and its role is to reduce the potential errors of limitations of the more simple / traditional methods.
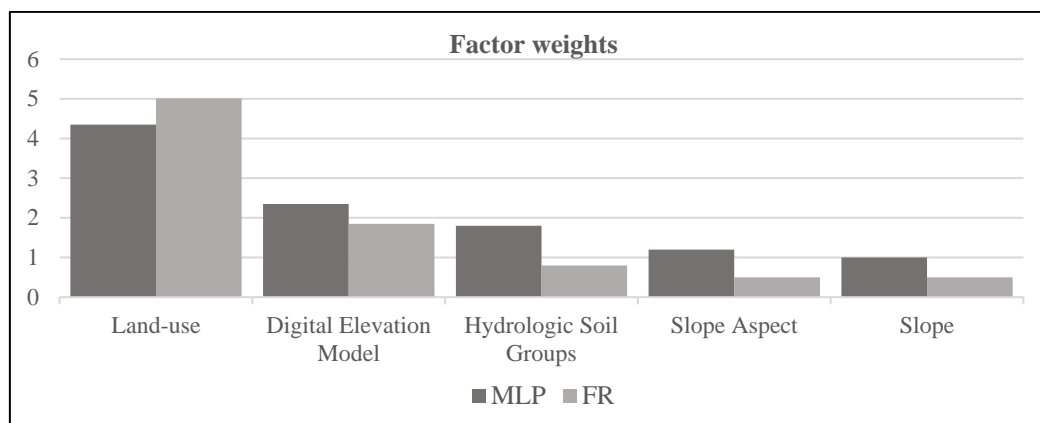
Figure 8. Geographical factor weights for the MLP and FR

RESULTS

After all the models proposed by this study were computed, the results obtained in a raster format have been converted from raster to vector (polygons) and the areas with a high potential have been aggregated and based on the distance from the river, land-use and the area covered, they were split into two categories: Flow control area and Permanent wetland. The results obtained in a raster format were classified into two classes, High potential areas and Low potential areas. Based on the results converted into vector format, the areas of the proposed wetlands were extracted (Figure 4). The vector results represent the area occupied by water within the proposed areas. The results show that the Frequency Ratio statistical model has the highest degree of coverage for both types of proposed areas, whilst the MLP-FR hybrid ensemble shows the lowest coverage rates. The hybrid model between the MLP and FR shows similar results to the that of the MLP neural network, with a difference of 2,09% for the permanent wetlands and 4,05% for the areas used for flow control. The surfaces resulting from the 3 models computed in the present study are overlapping agricultural lands, encoded as non-irrigated arable land in the CORINE Land Cover data set.
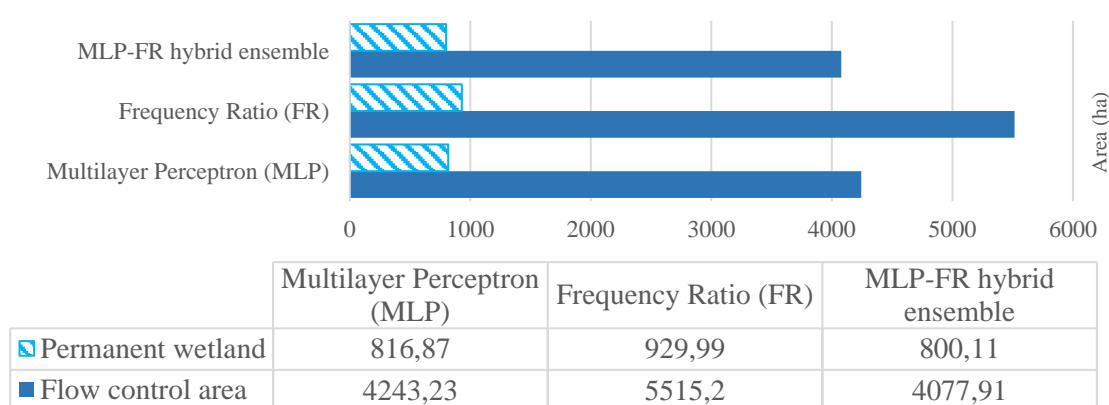

	Multilayer Perceptron (MLP)	Frequency Ratio (FR)	MLP-FR hybrid ensemble
Permanent wetland	816,87	929,99	800,11
Flow control area	4243,23	5515,2	4077,91

Figure 9. Mapped area total surface in hectares by the type of wetland
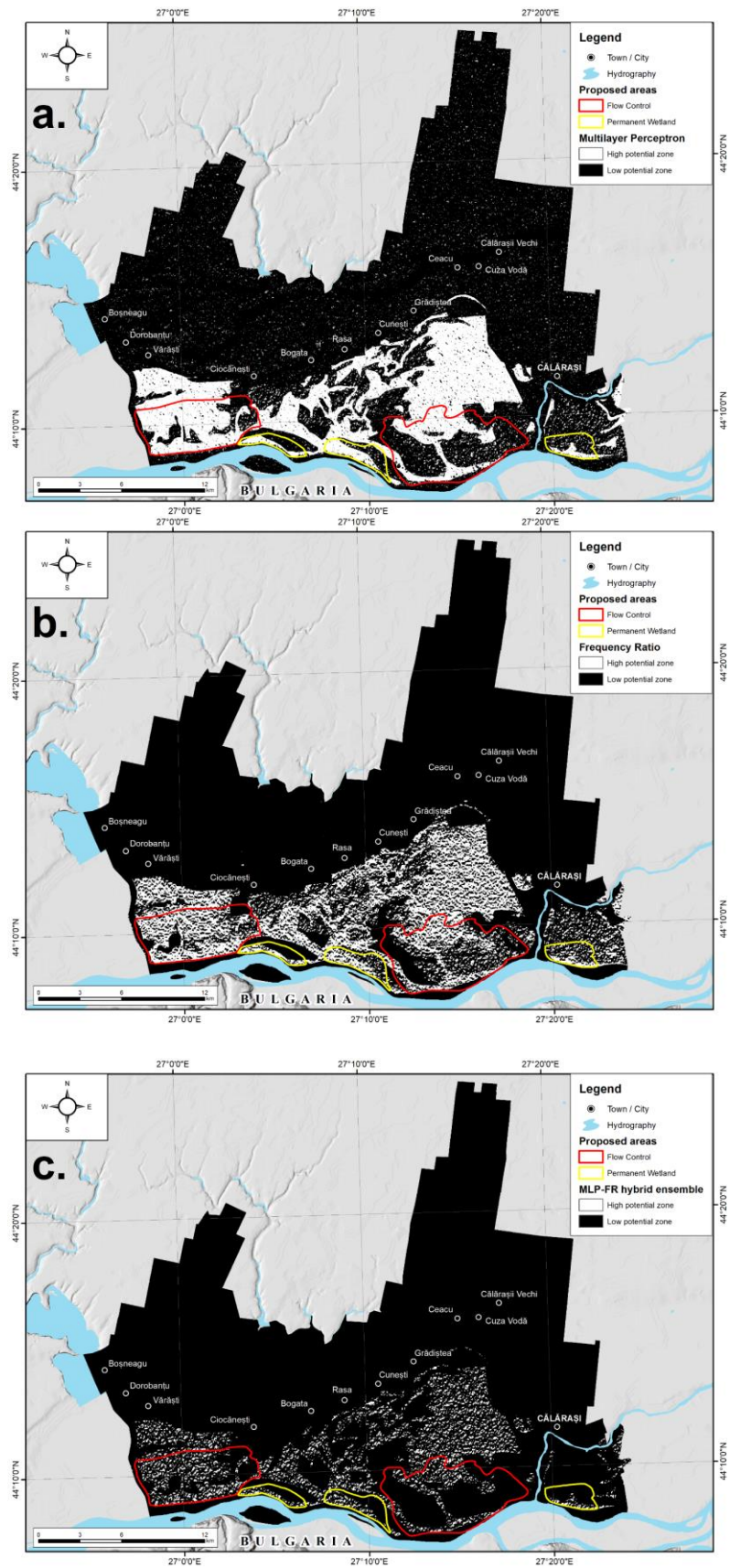
Figure 10. Initial results (a) Multilayer Perceptron, (b) Frequency Ratio, (c) MLP-FR hybrid ensemble
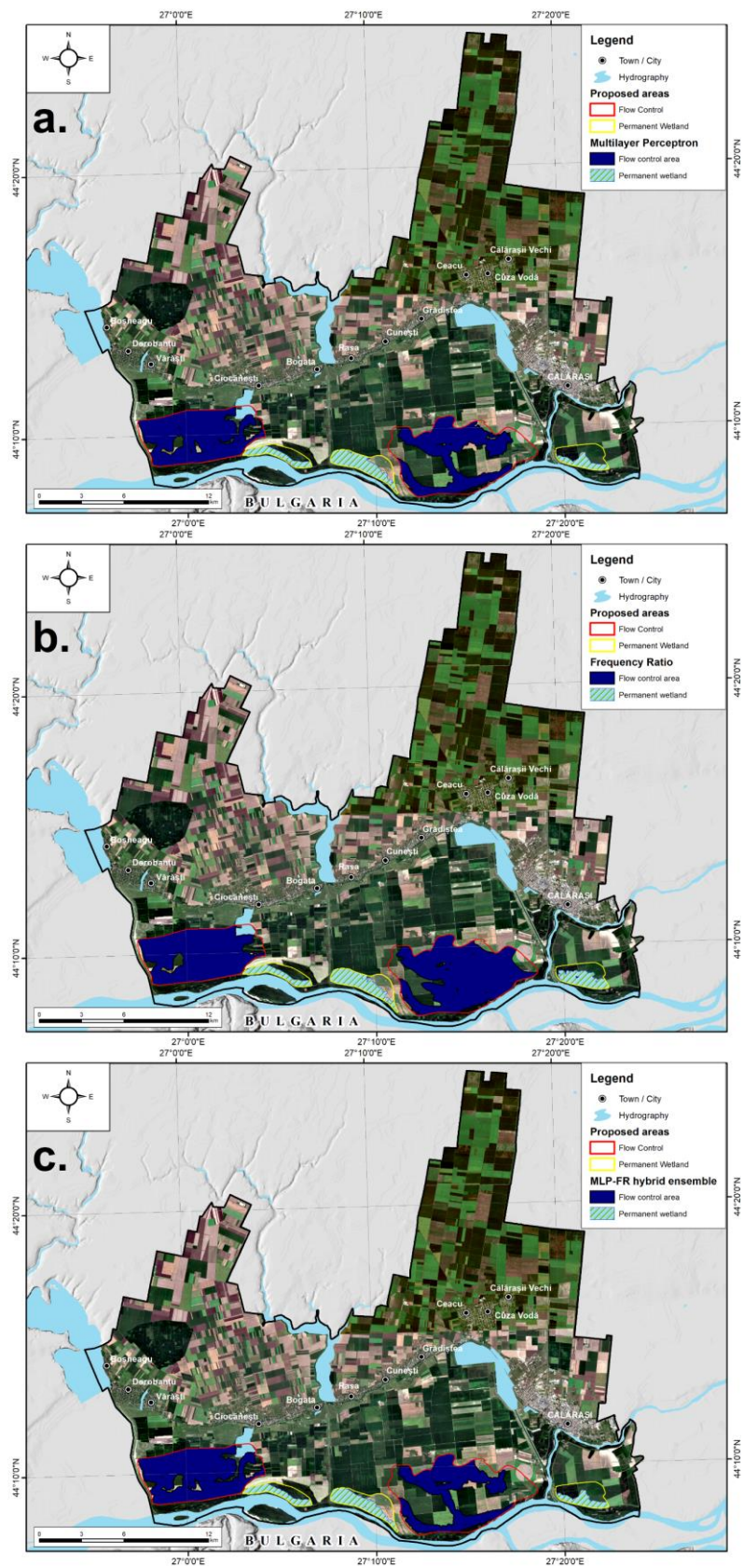
Figure 11. Final results (a) Multilayer Perceptron, (b) Frequency Ratio, (c) MLP-FR hybrid ensemble

CONCLUSIONS

The solutions presented in the study propose the creation of wetlands with a permanent or semi-permanent regime which have the purpose of storing and controlling water during high flows. The restoration and through creating these wetlands would help the development of new ecosystems thus leading to the creation of new spaces for fish spawning and the improvement of the flora. These new spaces could also serve recreational and tourist purposes. The creation of areas contributes permanently to the resolution of a number of imbalances identified in the number of previous studies [14-20].

The methodology developed in this study was applied on the lower Danube floodplain, between Boianu and Călărași, areas heavily affected during the extreme floods of 2006 and 2010. The proposal to improve this sector would lead to a better protection of the local communities against floods and would also lead to an improvement of the lateral connectivity of the river with its floodplain, contributing to the restoration of the local ecosystems. These new wetlands with a permanent or semi-permanent regime could also be used in the tourism sector or for creating fish farms. The role of these wetlands is to store and control the water during high flows that would normally not be supported by the riverbed. However, great efforts will be needed to create these wetlands, and as a result the areas covered by agricultural lands would be diminished permanently or temporarily. The benefits of creating these wetlands would offset the losses of agricultural land by reducing the losses caused to local communities, whether human or economic. The role of the areas used for flow control would ensure the agricultural lands a greater exchange of organic matter, which would to an increase of productivity, thus making the fertilization with chemical substances unnecessary.

REFERENCES

[1] Bularda M., Vişinescu I., 'Incintele îndiguite la Dunăre – Probleme actuale şi de perspectivă', Academia de Științe Agricole şi Silvice "Gheorghe Ionescu-Sisești", Stațiunea de Cercetare-Dezvoltare Agricolă Brăila, 2014.

[2] Țuchiu E. Gâştescu Petre, "The Danube river in the lower sector in two hydrological hypostases – High and Low waters," *Riscuri şi catastrofe*, vol. XI, no. (10,1), pp. 165–182, 2012.

[3] I. Mihnea, R. Mândru, and M. Bran, "Dams — necessity or calamity," *Carpathian Journal of Earth and Environmental Sciences*, vol. 3, no. 1, pp. 31–38, 2008.

[4] I. Nichersu and I. Nichersu, "Cercetări pentru studiul modificărilor hidromorfografice în lunca şi Delta Dunării, Buletinul AGIR, Supliment 3/2016," *Asociația Generală a Inginerilor din România*, 2016.

[5] D. C. Diaconu, "Particularities of Drain Liquid in the Small Wetland of Braila Natural Park, Romania," in *Water Resources Management in Romania*, A. M. Negm, G. Romanescu, and M. Zeleňáková, Eds. Cham: Springer International Publishing, 2020, pp. 437–464.

[6] WWF, "Proiectul PES implementat în bazinul Dunării » Zone pilot."

[7] R. Prăvălie, I. Sirodoev, C. Patriche, G. Bandoc, and D. Peptenatu, "The analysis of the relationship between climatic water deficit and corn agricultural productivity in the Dobrogea plateau," *Carpathian journal of earth and environmental sciences*, vol. 9, pp. 201–214, Nov. 2014.

[8] R. Prăvălie, C. V. Patriche, I. Sîrodoev, G. Bandoc, M. Dumitraşcu, and D. Peptenatu, "Water deficit and corn productivity during the post-socialist period. Case study: Southern Oltenia

drylands, Romania," *Arid Land Research and Management*, vol. 30, no. 3, pp. 239–257, Jul. 2016, doi: 10.1080/15324982.2015.1091399.

[9] I. Andronache *et al.*, "Assessment of Textural Differentiations in Forest Resources in Romania Using Fractal Analysis," *Forests*, vol. 8, no. 3, p. 54, Mar. 2017, doi: 10.3390/f8030054.

[10] I. C. Andronache *et al.*, "Fractal analysis for studying the evolution of forests," *Chaos, Solitons & Fractals*, vol. 91, pp. 310–318, Oct. 2016, doi: 10.1016/j.chaos.2016.06.013.

[11] M. C. Popa, D. Peptenatu, C. C. Drăghici, and D. C. Diaconu, "Flood Hazard Mapping Using the Flood and Flash-Flood Potential Index in the Buzău River Catchment, Romania," *Water*, vol. 11, no. 10, p. 2116, Oct. 2019, doi: 10.3390/w11102116.

[12] H. Gómez and T. Kavzoglu, "Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela," *Engineering Geology*, vol. 78, no. 1, pp. 11–27, Apr. 2005, doi: 10.1016/j.enggeo.2004.10.004.

[13] R. Sánchez-Reolid *et al.*, "Artificial Neural Networks to Assess Emotional States from Brain-Computer Interface," *Electronics*, vol. 7, no. 12, p. 384, Dec. 2018, doi: 10.3390/electronics7120384.

[14] A. Peponi, P. Morgado, and J. Trindade, "Combining Artificial Neural Networks and GIS Fundamentals for Coastal Erosion Prediction Modeling," *Sustainability*, vol. 11, no. 4, p. 975, Jan. 2019, doi: 10.3390/su11040975.

[15] "Multilayer perceptron architecture optimization using parallel computing techniques." [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0189369. [Accessed: 16-Jan-2020].

[16] M. F. Allawi *et al.*, "Reservoir Evaporation Prediction Modeling Based on Artificial Intelligence Methods," *Water*, vol. 11, no. 6, p. 1226, Jun. 2019, doi: 10.3390/w11061226.

[17] H. Khan, M. Shafique, M. A. Khan, M. A. Bacha, S. U. Shah, and C. Calligaris, "Landslide susceptibility assessment using Frequency Ratio, a case study of northern Pakistan," *The Egyptian Journal of Remote Sensing and Space Science*, vol. 22, no. 1, pp. 11–24, Apr. 2019, doi: 10.1016/j.ejrs.2018.03.004.

[18] H. Mojaddadi, B. Pradhan, H. Nampak, N. Ahmad, and A. H. bin Ghazali, "Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS," *Geomatics, Natural Hazards and Risk*, vol. 8, no. 2, pp. 1080–1102, Dec. 2017, doi: 10.1080/19475705.2017.1294113.

[19] E. Nohani *et al.*, "Landslide Susceptibility Mapping Using Different GIS-Based Bivariate Models," *Water*, vol. 11, no. 7, p. 1402, Jul. 2019, doi: 10.3390/w11071402.

[20] A. Yalcin, S. Reis, A. C. Aydinoglu, and T. Yomralioglu, "A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey," *CATENA*, vol. 85, no. 3, pp. 274–287, Jun. 2011, doi: 10.1016/j.catena.2011.01.014.