Scientific Papers. Series E. Land Reclamation, Earth Observation & Surveying, Environmental Engineering. Vol. XIII, 2024
Print ISSN 2285-6064, CD-ROM ISSN 2285-6072, Online ISSN 2393-5138, ISSN-L 2285-6064

ANALYTICAL FRAMEWORK ORIENTED TOWARDS ENVIRONMENTAL TRIGGERS IN EUROPEAN GREAT CITIES

Stefan-Mihai PETREA[1, 2, 4], Ira-Adeline SIMIONOV[1, 2, 5], Alina ANTACHE[1, 2, 6], Aurelia NICA[1], Maxim ARSENI[2, 3], Adrian ROSU[2, 3], Dragos CRISTEA[4], Catalina ITICESCU[2, 3]

[1]"Dunarea de Jos" University of Galati, Food Science and Engineering Faculty, 111 Domneasca Street, Galati, Romania
[2]"Dunarea de Jos" University of Galati, REXDAN Research Infrastructure, 98 George Cosbuc Street, Galati, Romania
[3]"Dunărea de Jos" University of Galati, Faculty of Sciences and Environment, 111 Domneasca Street, Galati, Romania
[4]"Dunărea de Jos" University of Galati, Faculty of Economics and Business Administration, 59-61 Nicolae Balcescu Street, Galati, Romania
[5]"Dunărea de Jos" University of Galati, Faculty of Automatic Control, Computers, Electrical Engineering and Electronics, 2 Stiintei Street, Galati, Romania
[6]"Alexandru Ioan Cuza" University of Iasi, 11 Carol I Blvd, Iasi, Romania

Corresponding author email: stefan.petrea@ugal.ro

Abstract

The paper aims to develop a comparative and integrated analytical framework (CIAF) that will ensure an in-depth understanding of environmental triggers in Western (W) vs Eastern (E) Great Cities (GC) of the European Union (EU). Parameters corresponding to both environmental (EVD) and economic (ECD) dimensions were selected for each of the first 10 EU WGC and EGC, respectively. The EVD considers the impact of PM2.5. exposure, as well as municipal waste generation and days of strong heat stress, while the ECD considers the GDP, labour productivity, as well as unemployment rate, all connected to the demographical dynamics of the analysed urban areas. A machine-learning methodology, consisting of MLR and XGBoost algorithms, is used for the development of the CIAF. The results indicate significant peculiarities between both WGC and EGC and reveal high accuracy (>85%) in various prediction scenarios. The findings can be used as a basis for the future development of complex decision-support tools, tackling to optimize the environmental management in EU GC.

***Key words**: environmental triggers, machine learning, great cities, PM2.5, municipal waste.*

INTRODUCTION

According to Cristea et al. (2022), environmental triggers and challenges extend their impact on economic and environmental dimensions. Also, Nuță et al. (2021) revealed that industrialization and intensive urbanisation can lead to environmental degradation, thus, affecting the environmental sustainability desideratum. However, the conclusion of the above-mentioned study reveals the hypothesis stating that urbanization may damage the environment only at the early stages, followed by beneficial effects further on.

A previous study (Cristea et al., 2022) emphasizes the peculiarities of various geopolitical blocks (BRICS vs. G7 vs. EU) and geographical regions related to environmental strategies and triggers. Rapid urban industrialization, transport agglomeration or intensive consumption behaviour can be considered human activity triggers, according to Simionov et al. (2019), that can alter environmental status in terms of water pollution with various contaminants of emerging concern (e.g. microplastics, pharmaceutical compounds) or priority pollutants (e.g. heavy metals) and can increase the air pollution (e.g. increase the concentrations of PM 2.5) according to Andrei et al. (2024), CO_2 emissions according to Nuță et al. (2021) and GHG according to Cristea et al. (2022). Also, Andrei et al. (2024) report a significant correlation between particulate pollution and welfare features in urban areas.

According to Petrea et al. (2022), the process of measuring sustainable development at an

aggregate level, requires a broad integration of indicators from various dimensions as economic, environmental, as well as social dimensions, since a symbiotic nexus between these dimensions is required to ensure the long-term development peculiarity.
Economic growth and technological development are revealed by indicators such as labour productivity and unemployment rate, while pollution intensity in urban areas can be estimated by considering air pollution and waste generation, considering the general background of stressors and pressures that create favourable conditions for climate change intensification.
Therefore, an integrated analytical framework that quantifies parameters from both economic and environmental dimensions, considering, at the same time, the climate change impact, can reveal innovative insides related to the conditionalities and relations between the parameters of both dimensions, including climate change.
However, considering that various regions have specific peculiarities possibly generated by their culture, political framework, habits, as well as industrial and technological development background, a comparative study between these regions, in order to identify various triggers that can have significant importance in environmental development strategy, is required.
Thus, the present study aims to develop a comparative and integrated analytical framework (CIAF), based on machine learning algorithms, that will ensure an in-depth understanding of environmental triggers in Western-Central (WC) vs Eastern (E) Great Cities (GC) of the European Union (EU), based the urbanization degree, considering various parameters related to both environment and economical dimensions.

MATERIALS AND METHODS

Data collection and dataset description

The dataset used for training and validation of the resulting models within the present study is structured, as follows:

- Environmental dimension:
  - PM2.5 [μg/m³], source: www.iqair.com;
  - municipal waste generation (MWG), source: Eurostat.
  - days of strong heat stress (DHS) [no. of days], source: OECD, Eurostat.
- Economic dimension
  - GDP, source: OECD, Eurostat.
  - labour productivity (LP), source: OECD, Eurostat.
  - unemployment rate (UR), source: OECD, Eurostat.

The dataset consists of 63 inputs, no null values, corresponding to both the WCGC and EGC groups. Each of the groups consists of 7 great cities, that were selected considering their geographical position, as well as their urbanization degree and data availability for a 9 years-period. Thus, the WCGC group covers data from Berlin (DE), Rome (IT), Paris (FR), Hamburg (DE), Munich (DE), Milan (IT), and Köln (DE), while the EGC group covers data from Warsaw (PL), Budapest (HU), Sofia (BUL), Krakow (PL), Zagreb (CRO), Wroclaw (PL), Lodz (PL).

Machine learning data-processing methods and workflow

The workflow used in the present study for data processing in order to develop the comparative and integrated analytical framework (CIAF) is presented in Figure 1. Thus, three machine learning-based supervised algorithms were used, namely the multiple linear regression (MLR), XGBoost (XGB) and Random Forest (RF), in order to generate high-metrics predictions for each of the analysed parameters, part of both dimensions. The workflow implies data preparation (a range between 70-80% of the data used for training the models, while the rest of 20-30% is used for validation purposes). This database pre-processing is described by Petrea et al. (2023), while the predictor's standardization was performed as presented by Petrea et al. (2020).

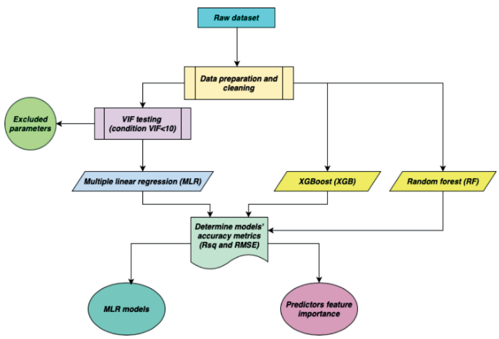

Figure 1. The workflow for CIAF development

Scientific Papers. Series E. Land Reclamation, Earth Observation & Surveying, Environmental Engineering. Vol. XIII, 2024
Print ISSN 2285-6064, CD-ROM ISSN 2285-6072, Online ISSN 2393-5138, ISSN-L 2285-6064

The MLR model equation is presented below (eq. 1), according to Petrea et al. (2023):

$$y_{predict} = a_1x_1 + a_2x_2 + \cdots + a_nx_n + b + \epsilon \quad (1)$$

where:
- ypredict is the dependent variable;
- $x_1, x_2 \ldots x_n$ are the n independent variables (predictors);
- b is the intercept indicating the *Y* value when all the predictors are zeros;
- $a_1, a_2 \ldots a_n$ are the coefficients of predictors, reflecting the contribution of each independent variable in predicting the dependent variable;
- ε is the residual term indicating the difference between the actual and the fitted response value.

According to previous publication (Chen and Guestrin, 2016), the XGBoost model has a core idea to fit and learn the residuals of the previous tree, with the final prediction result being the sum of the effects of all regression trees (eq. 2).

$$Obj = \sum_{i=1}^{n} l(y_i - \hat{y_i}) + \sum_{k=1}^{K} \Omega(f_k) \quad (2)$$

where:
- y_i is the measured value;
- $\hat{y_i}$ is the predicted value of each tree; l is the loss function, which is used to measure the total prediction error;
- $\sum\Omega(fk)$ is the regularization term.

The Random Forest algorithm is presented in eq. 3.

$$RFfi_i = \frac{\sum_{j \in all\ trees} normfi_{ij}}{T} \quad (3)$$

where:
- *T* is the total number of trees;
- $normfi_{ij}$ is the normalized future importance for i in tree j;
- $RFfi_i$ is the feature *i* importance, calculated from all the trees in the RF model.

RESULTS AND DISCUSSIONS

MLR models

Considering the WCGC group, it can be observed that the MLR model for predicting the DHS, considering only the environmental dimension, records values of VIF that do not exceed 10, meaning that all parameters are considered suitable to be used as predictors (Figure 2). The accuracy metrics of the model, presented in eq. 4 record the following values: R-sq 0.54, R-adj 0.53, RMSE 17.30 (Figure 3). It can be concluded that PM 2.5 parameter has the highest influence when predicting the DHS in the case of WCGC. Thus, intensive industrial activity and urbanization, considered triggers of PM2.5. in the great cities, have an immediate impact on climate change, manifested by the increase in the number of days that record temperatures of over 32ºC.

W-C DHS = -14.39 – 0.001 W-C EU - MWG + 3.55 W-C EU PM2.5 (4)

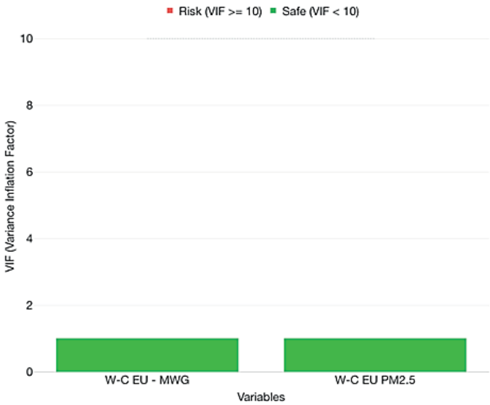

Figure 2. VIF values for the predictors, considering the environmental parameters for W-C DHS prediction

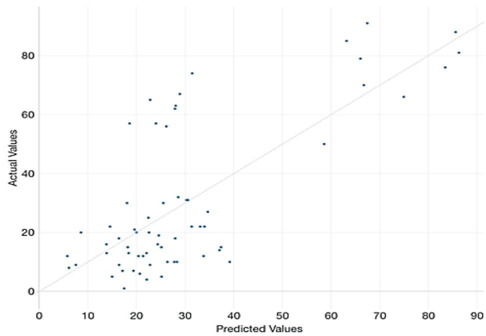

Figure 3. Actual/predicted value for W-C DHS prediction considering the environmental dimension

In the case of EGC group the prediction of DHS model revealed not significant ($p>0.05$) importance. Thus, the resulting model was considered unfit for the CIAF.

However, the use of economic dimension parameters as predictors in generating the DHS

Scientific Papers. Series E. Land Reclamation, Earth Observation & Surveying, Environmental Engineering. Vol. XIII, 2024
Print ISSN 2285-6064, CD-ROM ISSN 2285-6072, Online ISSN 2393-5138, ISSN-L 2285-6064

MLR model, for both WCGC and EGC, is doable. Therefore, the W-C DHS model is presented in eq. 5, while the E DHS model is presented in eq. 6. It can be observed that UR has the main influence in predicting the DHS, a fact valid for both WCGC and EGC. However, the decrease of UR in the WCGC generates an increase in DHS, a fact that may be due to the integration of extra labour in industry and, therefore, the increase of industrial production. The situation is reversed in the case of EGC, emphasizing that although the UR increased, the DHS still records an upward trend, a fact that may be due to the lack of industrial capacity, the labor market being oriented towards other niches with less impact on the environment.

$$W\text{-}C\ DHS = 607.14 - 0.007\ W\text{-}C\ EU\ GDP - 2.907\ W\text{-}C\ EU\ LP - 4.536\ W\text{-}C\ EU - UR \quad (5)$$

$$E\ DHS = 42.69 + 0.002\ E\ EU\ GDP - 0.407\ E\ EU\ LP + 1.628\ E\ EU - UR \quad (6)$$

The accuracy metrics of the models, presented in eq. 5 and 6 record the following values: R-sq 0.80, R-adj 0.79, RMSE 11.46 for WCGC (Figure 4) and R-sq 0.56, R-adj 0.53, RMSE 12.01 for EGC (Figure 5).

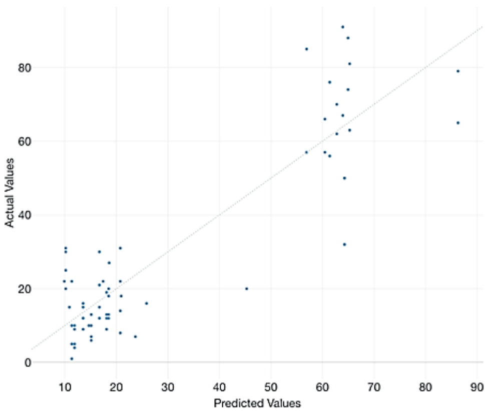

Figure 4. Actual/predicted value for W-C DHS prediction considering the economic dimension

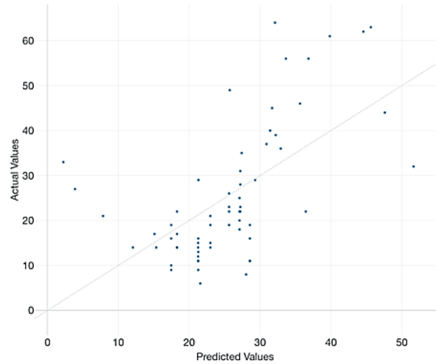

Figure 5. Actual/predicted value for E DHS prediction considering the economic dimension

All predictors were statistically significant ($p<0.05$) and all VIF values were under 2 (Figure 6) in the case of EGC and under 9 (Figures 7) in the case of WCGC.

When considering the W-C DHS prediction in relation to both economic and environmental parameters (Figure 8), the W-C UR and W-C GDP parameters record VIF values >10 – thus, these parameters are removed from the framework and the model's training and validation start considering the new formula of predictors. The resulted model is emphasized in eq. 7.

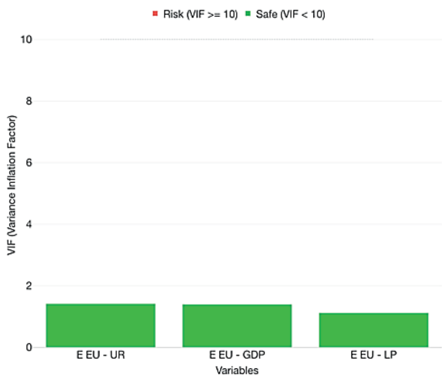

Figure 6. VIF values for the predictors, considering the economic parameters for E DHS prediction

Figure 7. VIF values for the predictors, considering the economic parameters for W-C DHS prediction

Figure 8. VIF values for the predictors, considering both the economic and environmental parameters for W-C DHS prediction

W-C DHS = 614.21 – 5.79 W-C EU LP – 0.004 W-C EU MWG + 1.72 W-C EU PM2.5. (7)

The VIF value related to eq. 7 model records values under 2. The accuracy parameters record the following metrics: R-sq 0.80, R-adj 0.79, RMSE 11.38 (Figure 9). All predictors were statistically significant ($p<0.05$).

The E DHS prediction considering both dimensions (economic and environmental) is emphasized through eq. 8. The values of VIF are under 2 (Figure 10) and all predictors were statistically significant ($p<0.05$). The model has inferior metrics compared to the WCGC correspondent model, as follows: R-sq 0.57, R-adj 0.52, RMSE 11.94 (Figure 11).

E DHS = 28.13 + 0.002 E EU GDP - 0.376 E EU LP + 0.008 E EU MWG + 1.877 E EU – UR + 0.027 E EU PM2.5 (8)

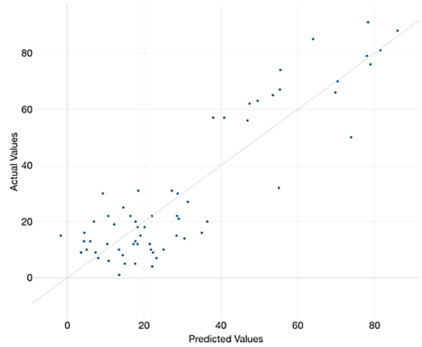

Figure 9. Actual/predicted value for W-C DHS prediction considering both the economic and environmental dimensions

Figure 10. VIF values for the predictors, considering both the economic and environmental parameters for E DHS prediction

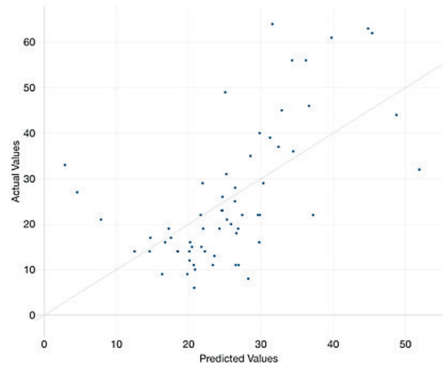

Figure 11. Actual/predicted value for E DHS prediction considering both the economic and environmental dimensions

The PM2.5 prediction, considering the environmental dimension predictors, reveal a not significant model ($p>0.05$) in the case of EGC, while for WCGC, the model (eq. 9)

Scientific Papers. Series E. Land Reclamation, Earth Observation & Surveying, Environmental Engineering. Vol. XIII, 2024
Print ISSN 2285-6064, CD-ROM ISSN 2285-6072, Online ISSN 2393-5138, ISSN-L 2285-6064

confirms the relation between PM2.5 and DHS. The model metrics are as follows: R-sq 0.54, R-adj 0.52, RMSE 3.58 (Figure 12). The VIF values were lower than 10.

W-C EU PM2.5. = 8.438 + 0.152 *W-C DHS* + *0.00007 W-C EU -MWG* (9)

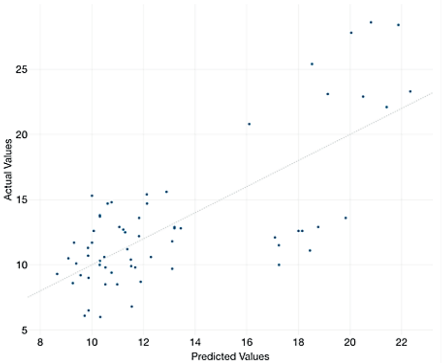

Figure 12. Actual/predicted value for W-C PM 2.5 prediction considering the environmental dimensions

The PM2.5. prediction models considering the economic dimension parameters as predictors were not significant in terms of importance ($p>0.05$). Thus, this reveals the incapacity of using economic dimension in predicting PM2.5 - however, the number of dataset inputs can be considered a limitation barrier in the process of identifying high accuracy and significant prediction models, considering the above-mentioned background. This situation repeats in when considering the parameters of both dimensions for developing PM2.5 prediction models for EGC and WCGC, as well as in the scenario of developing MWG prediction models considering the environmental dimension parameters as predictors.

The MLR prediction models for MWG that consider the economic dimension parameters as predictors are doable, both in the case of WCGC (eq. 10) and EGC (eq. 11) and emphasize the importance of UR in predicting MWG.

W-C EU MWG = -15.76 + 0.638 W-C EU GDP – 79.89 W-C EU – LP +753.859 W-C EU-UR (10)

E EU MWG = 1630.59 - 0.047 E EU GDP – 3.641 E EU – LP - 28.547 E EU-UR. (11)

All predictors were statistically significant ($p<0.05$), except LP in the case of W-C MWG model. All VIF values were under 2 (Figure 13) in the case of EGC and under 9 (Figure 14) in the case of WCGC. The models' accuracy metrics are as follows: Rsq 0.622, Radj 0.600, RMSE 780.26 for W-C MWG and Rsq 0.55, Radj 0.50, RMSE 157.70 for E MWG.

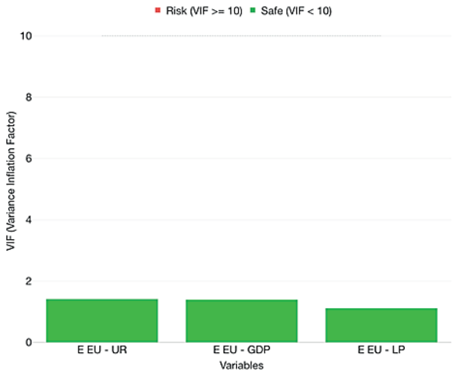

Figure 13. VIF values for the predictors, considering the economic parameters for E MWG prediction

Figure 14. VIF values for the predictors, considering the economic parameters for W-C MWG prediction

If considering all parameters from both dimensions for generating MLR prediction models for W-C MWG and E MWG, it can be observed that in the case of W-C MWG the GDP predictor records a VIF value of 16 (over the limit of 10) and is excluded from the prediction framework. Thus, in terms of statistical significance, in the case of WC MWG the LP and PM2.5 predictors are not significant ($p>0.05$), while the rest of the predictors are statistically significant ($p<0.05$). Also, in the case of predicting the E MWG, the predictors

DHS, LP and PM2.5 are statistically not significant ($p>0.05$), while GDP and UR and statistically significant ($p<0.05$). The VIF values are under 4 in the case of WC MWG model (Figure 15), while for E MWG all the VIFs are under 2 (Figure16).

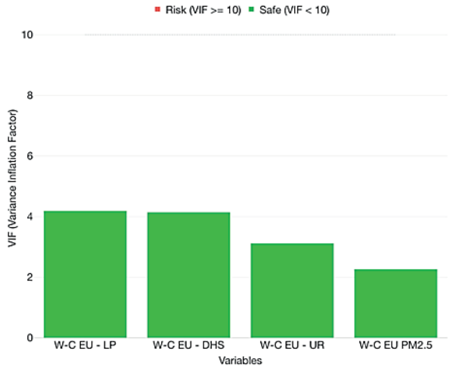

Figure 15. VIF values for the predictors, considering both the economic and environmental parameters for W-C MWG prediction

Figure 16. VIF values for the predictors, considering both the economic and environmental parameters for E-MWG prediction

The prediction of MWG considering all predictors from both dimensions (economic and environmental) reveals LP and UR as main predictors in the case of both WCGC and EGC groups, having the highest weight in predicting the dependent variable. The MWG and LP are in an indirect relation, a fact valid for both groups, while in relation to UR, the predicted variable presents a direct relation for WCGC and an indirect relation for EGC, a situation confirmed also by previous models (eq. 12, 13).

W-C EU MWG=11 867.36 – 44.976 W-C EU DHS – 107.036 W-C EU – LP + 327.495 W-C EU UR – 11.898 W-C EU PM2.5 (12)

E EU MWG = 1583.27 + 1.480 E EU-DHS – 0.050 E EU – GDP – 3.010 E EU – LP – 0.809 E EU – PM2.5 – 31.05 E EU – UR (13)

The models' accuracy metrics are as follows: Rsq 0.513, Radj 0.480, RMSE 885.77 for W-C MWG and Rsq 0.54, Radj 0.50, RMSE 156.64 for E MWG.

The RF and XGB models

The RF and XGB models were applied in predicting each of the environmental dimension parameters, considering the rest of the parameters from both economic and environmental dimensions. A comparative analysis was performed in order to establish which of the algorithms performs better, considering the accuracy metrics associated to the resulted models.

Thus, the XGB model for predicting W-C DHS presents high metrics (Rsq 0.988, RMSE 2.79) and reveals that GDP has the highest feature importance, compared to the rest of the predictors (Figure 17).

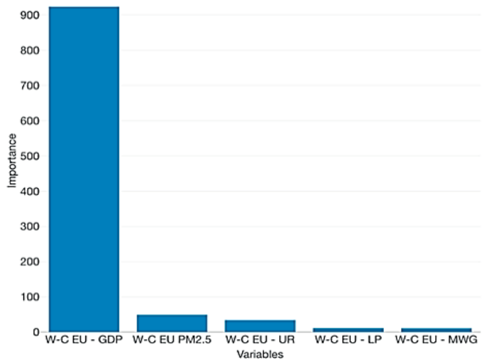

Figure 17. The predictors feature importance when predicting the W-C DHS using the XGB algorithm

This emphasizes the possible impact of economic growth on environmental status, within the WCGC. The RF model reveals that besides GDP, the LP has the highest feature importance (Figure 18), emphasizing a possible conditionality of the economic growth - environmental status relation by the intensity of the economic activities and, thus, the use of state-of-the-art technologies.

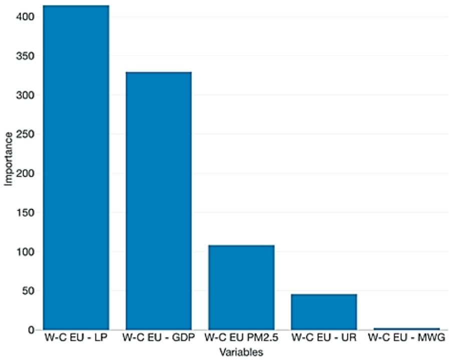

Figure 18. The predictors feature importance when predicting the W-C DHS using the RF algorithm

However, the RF reveals lower accuracy metrics (Rsq 0.854, RMSE 9.79), compared to XGB. The XGB model for predicting E DHS presents high metrics (Rsq 0.943, RMSE 44.91) and reveals that UR has the highest feature importance, compared to the rest of the predictors (Figure 19), confirming the results recorded in the MLR models.

This emphasises the low conditionality between the UR and GDP within the EGC, considering, simultaneously, the impact on the environment.

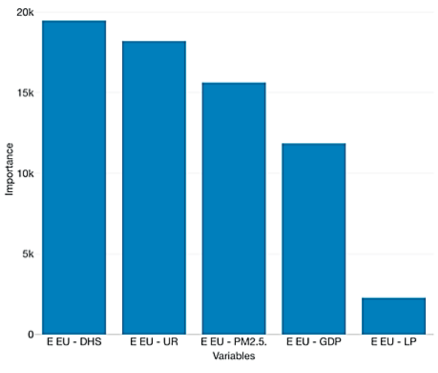

Figure 19. The predictors feature importance when predicting the E DHS using the XGB algorithm

The RF model reveals that E DHS is mostly impacted by LP and, therefore, technological development, in the case of EGC (Figure 20). However, the RF reveals low accuracy metrics (Rsq 0.58, RMSE 10.88).

The XGB model for predicting W-C PM2.5 presents high metrics (Rsq 0.956, RMSE 1.11) and reveals that DHS and MWG have the highest feature importance, compared to the rest of the predictors (Figure 21). This emphasises the possible impact of climate change and waste generation on PM 2.5. The RF model reveals that besides DHS, the LP has the highest feature importance (Figure 22), thus, confirming the conditionality of the economic growth on environmental status. However, the RF reveals lower accuracy metrics (Rsq 0.609, RMSE 3.31), compared to the XGBoost W-C PM2.5 predictability model.

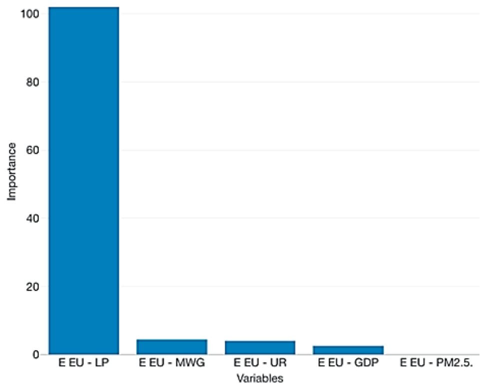

Figure 20. The predictors feature importance when predicting the E DHS using the RF algorithm

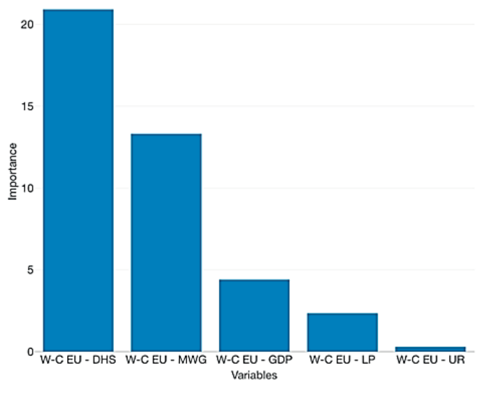

Figure 21. The predictors feature importance when predicting the W-C PM2.5 using the XGBoost algorithm

The XGBoost model for E PM2.5 (Rsq 0.853, RMSE 2.08) emphasizes that MWG has the highest feature importance, followed by DHS and GDP, the fact that emphasizes a peculiar problem related to waste management in the eastern bloc, a situation which could be conditioned by the lack of proper legislation or control in the area of environmental protection (Figure 23).

The E PM2.5 RF model reveals a completely different feature importance ranking among predictors, compared to W-C PM2.5 (Figure 24). Thus, it can be observed (Figure 24) that the first two predictors in terms of feature importance are from the economic dimension (GDP and UR), revealing the conditionality of the economic activities and economic development rate on environmental status and indicating the need to maintain sustainable economic development, limiting the impact on the environment and creating resilience to climate change. However, the RF reveals lower accuracy metrics (Rsq 0.609, RMSE 3.31), compared to the XGBoost W-C PM2.5 predictability model.

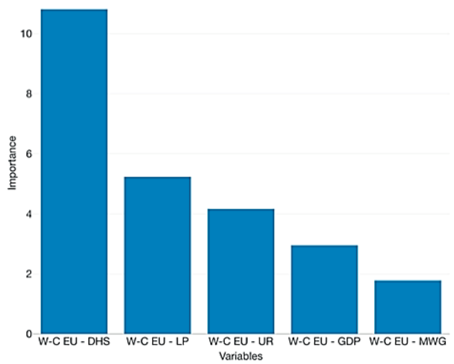

Figure 22. The predictors feature importance when predicting the W-C PM2.5 using the RF algorithm

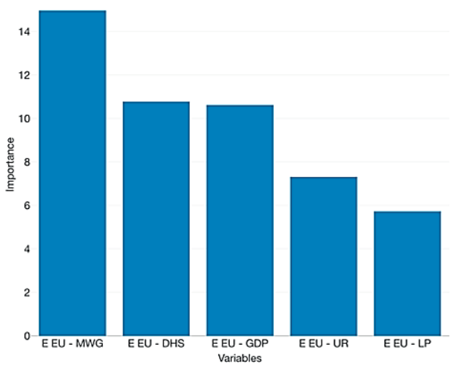

Figure 23. The predictors feature importance when predicting the E PM2.5 using the XGBoost algorithm

The XGB model for predicting W-C MWG presents the best metrics of accuracy from the present study analytical framework (Rsq 0.99, RMSE 102.24) and reveals the significant impact of GDP as the main predictor of this model, with the highest feature importance (Figure 25).

The RF model confirms the XGBoost findings (Figure 26) - however, the accuracy metrics are lower (Rsq 0.75, RMSE 627.02), compared to XGB.

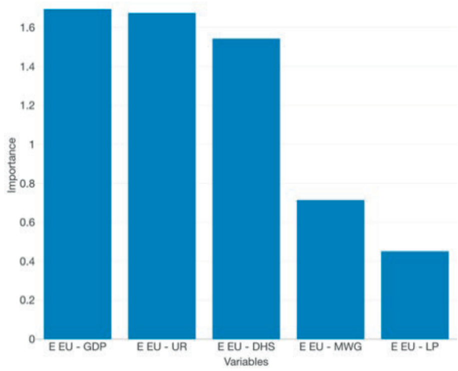

Figure 24. The predictors feature importance when predicting the E PM2.5 using the RF algorithm

The XGB model for predicting E MWG reveals a different framework compared to the one associated with the W-C block (Figure 27). Thus, DHS and UR are the main predictors in terms of the feature importance value. The RF does not perform well in predicting E MWG (Rsq 0.41, RMSE 145.60) (Figure 28). Thus, even if GDP and PM2.5 are found as predictors with the highest feature importance in determining the E - MWG concentration, their reliability is poor.

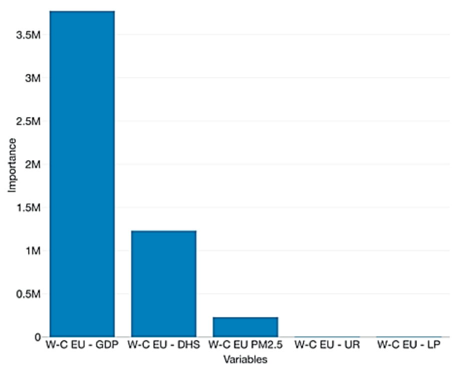

Figure 25. The predictors feature importance when predicting the W-C MWG using the XGBoost algorithm

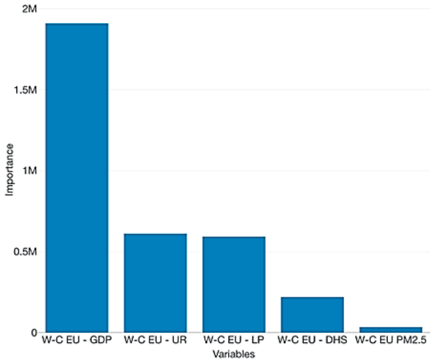

Figure 26. The predictors feature importance when predicting the W-C MWG using the RF algorithm

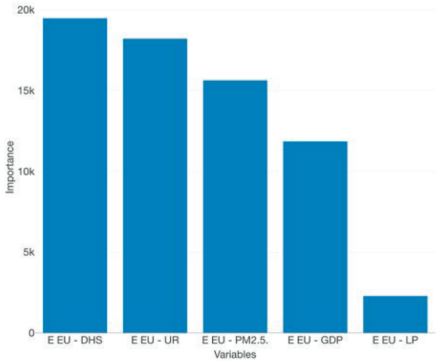

Figure 27. The predictors feature importance when predicting the E MWG using the XGBoost algorithm

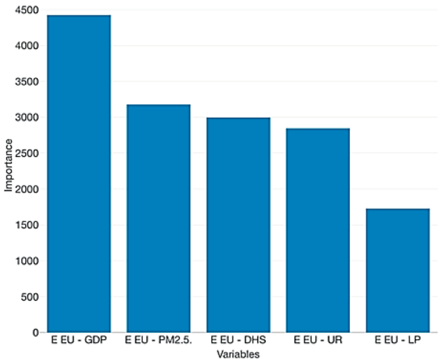

Figure 28. The predictors feature importance when predicting the E MWG using the RF algorithm

Similar to other studies, the present study has also limitations. Thus, firstly, the dataset was conditioned by the existing data – therefore, various GCs from both W-C and E blocks were not included due to the unavailable data. Also, the dataset covers a limited period and limited parameters that define both dimensions. It is recommended for future studies to extend the dataset to integrate other parameters and, also, to extend the number of dimensions, in order to achieve better ranking in terms of the integrative approach and better understanding in terms of triggers that impact the environmental status.

CONCLUSIONS

The results indicate significant peculiarities between both WGC and EGC and reveal high accuracy (>85%) in various prediction scenarios. The findings can be used as a basis for the future development of complex decision-support tools, tackling to optimize the environmental management in EU GC. The WCGC analytical framework reveals better predictability compared to the EGC.

There is a significant conditionality between the environmental and economic dimensions in both W-C and E GCs. However, it seems that the economic structure and development represent decisive triggers that impact the environmental status. All AI algorithms used for the present study reveal high accuracy metrics – however, it seems that XGB is associated with the best metrics, followed by MLR and RF.

ACKNOWLEDGEMENTS

The present research was supported by the project: "An Integrated System for the Complex Environmental Research and Monitoring in the Danube River Area", REXDAN, SMIS code 127065, co-financed by the European Regional Development Fund through the Competitiveness Operational Programme 2014-2020, contract no. 309/10.07.2021. The present research was supported by the project". Integrated research and sustainable solutions to protect and restore Lower Danube Basin and coastal Black Sea ecosystems, (ResPonSE)", 760010/30.12.2022, component C9. PRIVATE SECTOR SUPPORT, RESEARCH, DEVELOPMENT AND INNOVATION, Investment "I5. Establishment and operationalization of Competence Centers". (specific project no.1 - Decision Support Solutions based on mathematical modeling, for complex ecological systems – CESMoSS).

REFERENCES

Andrei, J. V., Cristea, D. S., Nuţă, F. M., Petrea, Șt. M., Nuţă, A. C., Tiron-Tudor, A. & Chivu, L. (2024). Exploring the causal relationship between PM 2.5 air pollution and urban areas economic welfare and social-wellbeing: Evidence form 15 European capitals, *Sustainable Development – Wiley*, https://doi.org/10.1002/sd.2920.

Chen, T., Guestrin, C. & Boost, X.G. (2016). A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*; Association for Computing Machinery: New York, NY, USA, pp. 785–794.

Cristea, D.S., Zamfir, C.G., Simionov, I.A., Fortea, C., Ionescu, R.V., Zlati, M.L., Antohi, V.M., Munteanu, D. & Petrea, S.M. (2022). Renewable energy strategy analysis in relation to environmental pollution for BRICS, G7, and EU countries by using a machine learning framework and panel data analysis. *Front. Environ. Sci., 10,* 1005806. doi: 10.3389/fenvs.2022.1005806.

Nuţă, F.M., Nuţă, A.C., Zamfir, C.G., Petrea, S.M., Munteanu, D., & Cristea, D.S. (2021). National Carbon Accounting-Analyzing the Impact of Urbanization and Energy-Related Factors upon CO_2 Emissions in Central–Eastern European Countries by Using Machine Learning Algorithms and Panel Data Analysis. *Energies, 14*, 2775. https://doi.org/10.3390/en14102775

Petrea, Ș.M., Cristea, D.S.; Rahoveanu, T.; Zamfir M.M., C.G., Zugravu A., Nancu G.A. (2020). Perspectives of the Moldavian Agricultural Sector by Using a Custom-Developed Analytical Framework. *Sustainability, 12*, 4671. https://doi.org/10.3390/su12114671

Petrea, Ș. -M., Simionov, I. A., Antache, A., Cristea D. S., Nica, A. & Cristea V. (2023). Analytical framework for the development of water quality virtual sensors in integrated aquaculture systems based on various aquaponics techniques. *IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA)*, Sinaia, Romania, pp. 1-5. doi: 10.1109/ETFA54631.2023.10275350.

Petrea, Ș.M., Costache, M., Cristea, D.S., Strungaru, A., Simionov, I.A., Mogodan, A. et al. (2020). A Machine Learning Approach in Analyzing Bioaccumulation of Heavy Metals in Turbot Tissues. *Molecules*, *25*, 4696.

Simionov, I.A., Cristea, V., Petrea, S.M., Bocioc Sîrbu, E. (2019). Evaluation of heavy metals concentration dynamics in fish from the Black Sea Coastal area: an overview. *Environmental Engineering and Management Journal*, 18, 5, 1097-1110 http://www.eemj.icpm.tuiasi.ro/; http://www.eemj.eu.