Assessing the Relative Importance of the Drivers of CO_2 Emissions in the Selected Emerging Economies Using Machine Learning Approach

Seema Joshi[1], Sachin Gupta[2*], Charu Kaistha[3]

[1]Department of Commerce, Kirori Mal College, University of Delhi, Delhi, India
[2*]Vivekananda School of Economics, Vivekananda Institute of Professional Studies- Technical Campus, Delhi, India
[3]Senior General Manager, Power Finance Corporation, Delhi, India
[2*]sachin.gupta@vips.edu

Abstract. The main objective of the present research is to answer a key research question: what is the relative importance of the drivers of CO2 emissions? Another important question the present study addresses is how the countries are related to each other regarding CO2 emissions. Taking a sample of 42 emerging economies from Asia and Sub-Saharan Africa (SSA) and using hierarchical clustering and the neural network method the study tries to answer the key question. Firstly, the explanatory variables were identified through a review of the literature. Subsequently, the gathered data was classified into two clusters having similar characteristic variables utilizing the dendrogram by performing an exploratory clustering method known as hierarchical clustering. Later using the machine learning K-Means clustering technique, the clusters were verified. The use of another machine learning method of feed-forward multilayer perceptron commonly known as neural network helped us to identify the relative importance of explanatory variables viz. economic growth (EG), renewable energy consumption (REC), urbanization (URB), democracy index (DI) and foreign direct investment (FDI) for their relation to the response variable viz. CO2 emissions. The neural network results reveal that EG, REC, and URB are the most important variables (with Rank 1,2, and 3 respectively) followed by DI (Rank 4) and FDI (Rank 5). FDI seemingly is the least important among these identified variables.

Keywords: CO2 emissions, economic growth (EG), renewable energy consumption (REC), urbanization (URB), democracy index (DI), and foreign direct investment (FDI).

1. Introduction

CO_2 emissions and its worsening impact on environment has preoccupied the attention of the academicians, activists, researchers and policy makers for the past few decades. As per one estimate, emerging economies account for almost two-thirds of the global GDP growth and more than half of new consumption over the past few years (MGI, 2018). These economies are guided by a pro-growth policy agenda. But economic growth and greenhouse gas emissions are closely tied up through most of modern economic history (The International Economic Agency i.e. IEA 2024). Besides, growing thrust on growth leads to engagement in anti-environmental activities which drive countries towards climate change and have detrimental impact on environmental quality (Shah et al., 2022). These developments necessitate that the drivers of CO_2 emissions should be identified, ranked in the order of their importance in influencing CO_2 emissions and dealt with to prevent the environmental damage. In one of the recent studies on the subject (Joshi et al. 2024), the economic and political determinants of CO_2 emissions have been identified. In the present research, the focus is on assessing the relative importance of the determinants of CO_2 emissions in the emerging economies of Asia and Sub-Saharan Africa.

Numerous studies have been devoted to the research on the determinants/drivers of CO_2 emissions globally (Glavas, 2024; Qiwen et al., 2021; Xia et al., 2020; Balogh and Jambor,2017). Besides, region specific (Ab-Rahim, R., Xin-Di,2021; Zmami and Ben Selha,2020; Morales-Lage et al., 2016; Iwata et al., 2012); country specific (Adebayo et al, 2021; 2012) and city specific studies (Zheng et al., 2019) too have attributed to this subject. However, studies focussing especially on the determinants of CO_2 emissions relating to emerging economies of Asia and Sub-Saharan Africa are hardly available.

The present research is relevant to the environmental pollution literature because the studies done thus far focused primarily on the determinants of environmental quality. As per our knowledge there is hardly any study which makes an attempt to assess and analyse the relative importance of the drivers of CO_2 emissions (proxy for environmental quality in our study) in the selected emerging economies from Asia and SSA. That is how the present paper is relevant as it tries to fill up these research gaps. We have identified five explanatory variables from the review of literature and gathered data on all these variables for 42 countries.

A. Jain and S. Gupta (eds.), *Proceedings of the International Conference on Policies, Processes and Practices for Transforming Underdeveloped Economies into Developed Economies (PPP-UD 2025)*, Advances in Economics, Business and Management Research 357,
https://doi.org/10.2991/978-94-6463-894-3_18

The first explanatory variable identified by us that affects CO_2 emissions or in other words environmental quality is economic growth (proxied by GDP per capita in our study). There is lot of literature that tries to debate the very existence of the U-shaped relationship between economic development and environmental degradation proposed by the Environmental Kuznets curve (EKC) (Grossman and Kruger,1991; Beckerman,1992; Arrow etal.,1995; Shafik,1994, Ahmad et al., 2017). In this study, the purpose is not to explore whether such a relationship exists or not. But we are interested in knowing the relative importance of this explanatory variable vis-a vis other explanatory variables in affecting CO_2 emissions.

The second explanatory variable included in this study is renewable energy (RENE). With the rise in energy demand globally, the bulk of future energy demands is likely to be met by various forms of RENE viz. solar, wind, ocean, hydro, biomass and geothermal etc. RENE offers itself as an alternative to fossil fuels (Moriarty and Honnery,2002; Joshi et al., 2024) which is a major energy source at present and accounted for 74% of all CO_2 emissions (Sims et al., 2007). The potential of RENE is still untapped and therefore calls for attention by the researchers as well as policy makers. As a source of clean energy, an alternative to fossil fuels, RENE can help to limit global warming and lead on a pathway to sustainable growth. The relative importance of renewable energy (RE) in causing carbon emissions is explored in the present paper.

Urbanization which is regarded as "one of the biggest social transformations of modern time, driving and driven by multiple social, economic, and environmental processes" (Bai et al. ,2017). Its profound, far-reaching effects and causal relationship with environment has been well captured in the literature (Wang; 2014; Wang et al., 2014; Wang et al. ,2016; Wang, Chen and Kubota,2016; Han et al., 2017; Zhang et al., 2017). This paper will try investigating for selected Asian and for SSA countries the role of urbanisation in impacting environmental quality. Undeniably, the role of energy availability is critical for economic development (Adams and Nasiah, 2019) but Surroop and Raghoo (2018) point out in their study that climate change and energy use are politically determined. The implementation of environmental energy policies largely depends on the varying incentives created by the political institutional environments (Bernauer and Koubi, 2009; Sarkodie and Adams, 2018). Therefore, it is essential to know the role played by political institutions and political system in impacting environmental quality. We are capturing the effect of political system by using the democracy index (DI) prepared by EIU. The effective governments can have a moderating effect on environment quality (Dulal et al., 2013, Joshi and Kansil,2024). This can happen via granting aids to businesses or by offering support in facilitating transition from a brown to a green economy.

Yet another explanatory variable used in this paper is foreign direct investment (FDI). It is generally viewed that globalisation and consequent expansion of economic activities results into environmental degradation (Huwart and Verdier, 2013). Since FDI contributes significantly to phenomenon called globalisation, therefore, concerns are raised sometimes about its environmental ramifications (Pazienza, 2019). There are conflicting hypotheses raised in the literature viz. the pollution halo and pollution haven hypotheses. As per the former hypothesis, FDI leads to green growth as environmentally friendly technologies are transferred by multinational companies to the host countries or in other words, this hypothesis argues for a positive relationship between FDI and environment (Tukhtamurodov et al., 2024 and Khan et al., 2023). Conversely, the latter hypothesis offers an alternative and negative view of FDI. It emphasizes that FDI may have environmental damaging effects. This may be due to lack of stringent environmental regulations in the recipient country/countries that are usually emerging ones (Copewell and Taylor, 1994 and Ketchova et al., 2024)

2. Methodology

The present study makes use of the cross-section data for the year 2020/2021 for 42 emerging economies from Asia and Sub-Saharan Africa (SSA). We have tries to assessing and analyse the relative importance of the chosen explanatory variables viz. economic growth (EG) proxied by Gross National Income per capita (GNIPC), renewable energy consumption i.e. REC (as a percentage of total final energy consumption), foreign direct investment i.e. FDI (as a percentage of GDP), urbanization (URB) and democracy index (DI) in causing or impacting the CO_2 emissions per capita. Therefore, CO_2 emissions per capita is the dependent variable in this study. It's important to mention here that higher CO_2 emissions are indicative of poor environmental quality and vice-versa. We have used the data from the World Bank, World Development Indicators (WDI), United Nations, United Nations Development Programme (UNDP), and the Economist Intelligence Unit Limited.

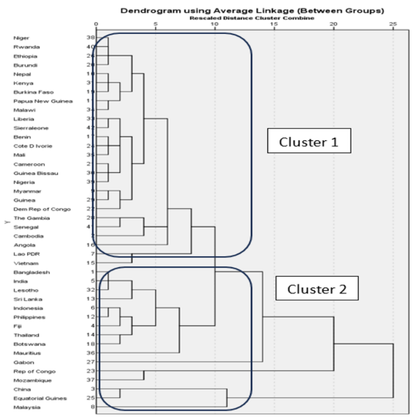
After gathering the data from diverse sources, the classification of the data was done into two clusters having similar characteristic variables utilizing the dendrogram by performing exploratory clustering method known as hierarchical clustering (Nielsen,2016). Subsequently, the machine learning K-Means clustering technique (Forgy, 1965) validated /verified the clusters. The validation of cluster is done by silhouette measure of cohesion and separation (Peter J. Rousseeuw, 1987)The use of another machine learning method of feed forward multilayer perceptron commonly known as neural network (Ojha et al., 2017) helped us to identify the relative importance of the explanatory variables viz. economic growth (EG), renewable energy consumption (REC), urbanization (URB), democracy index (DI) and foreign direct investment (FDI) for their relation to the response variable viz. CO_2 emissions.

3. Empirical Exercise

Keeping the key objective of the study in mind i.e. to assess the relative importance of identified explanatory variables in causing or contributing to CO_2 emissions/ environment quality, Multilayer perceptron neural network is built using SPSS and the importance score of each variable is calculated in causing CO_2 emissions.

After gathering the data, our first step was to identify the number of clusters present in the dataset. For this purpose, hierarchical clustering is used as an exploratory analysis. In this method of clustering, the clusters are identified on the basis of inter linkage distance. The less is the interlinkage distance among the entities more are the chances that they are depicting same behaviour and hence they are clubbed together in the same cluster. The squared Euclidean distance is used to measure proximity using average linkage. The dendrogram is shown in edited fig 1. The dendrogram clearly depicts us that the countries there can be partitioned into two clusters.

Fig 1: Hierarchical clustering



Source : Authors' calculations

Once the exploratory phase of the clustering is done, then there arises a need to identify the clusters using some conclusive method. K-mean clustering method helps us in this process of validation. Using K-means clustering we get the information about the cluster centers and cluster member countries out of a sample of 42 countries. As input from hierarchical clustering, K is chosen as two. It is also an effective method to identify the variables which are playing a role in clustering and how these variables are different in values with respect to the cluster. Initial cluster centers are shown in table 1, which shows that the cluster centers are having different values with respect to the clusters for variables. After identifying the initial centers which are selected on the random basis, an iterative algorithm is performed which computes distance from the center and assigns to the nearest cluster. The process is iterative in nature. Once there is no change in the cluster number the iteration stops.

Table 1: Initial Cluster Centers

	Cluster

CO2	8.40	.10
URB	77.20	13.70
FDI	5.00	.30
REC	5.80	83.50
DI	7.24	2.13
EG	26658	732

The parameter set in SPSS to perform at most 10 iterations, the convergence history tells us that the clusters converge within four iterations and hence the clusters are found to be stable shown in table 2.

Table 2: Iteration History

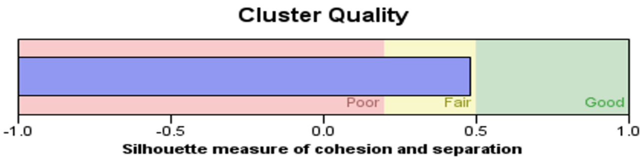
Iteration	**Change in Cluster Centers**	
	1	2
1	6775.02	3910.13
2	1974.49	476.03
3	1364.13	460.86
4	.000	.000

The membership of the countries after the clustering is performed is shown in the Table 3 with the cluster distance in appendix.

But the problem with K mean clustering solution is the ratio in which countries are divided. Here number of countries which are in cluster 1 is only 9 whereas cluster 2 is having 33 countries. The ratio of these numbers should not be more than 3, which violates in the present case as the ratio is more than 3 (=33/9).

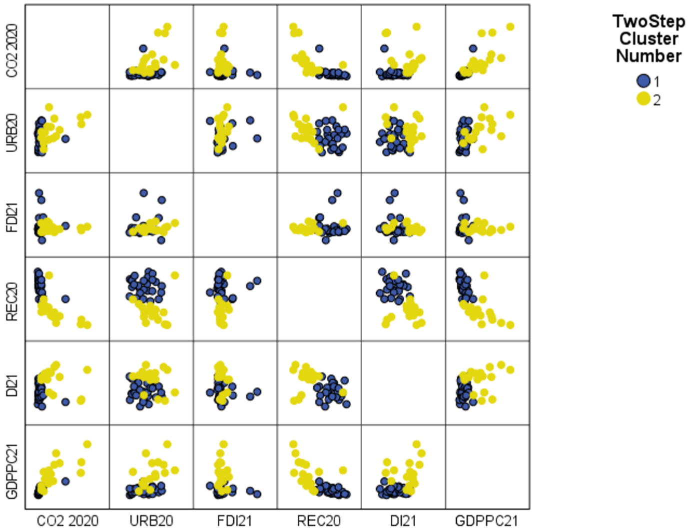
That's why two stage clustering is implemented which is based on silhouette measure of cohesion and separation. The five variables (economic growth (EG), renewable energy consumption (REC), urbanization (URB), democracy index (DI) and foreign direct investment (FDI)) are used as input to form a cluster and CO2 is put into the evaluation field. The result shows that the countries can be divided into clusters. One cluster is having 26 countries and another cluster is having 16 countries. This better way as the ratio is less than 2 (=26/16). The result is shown in figure 2.

Fig 2: Two stage clustering result based on silhouette measure



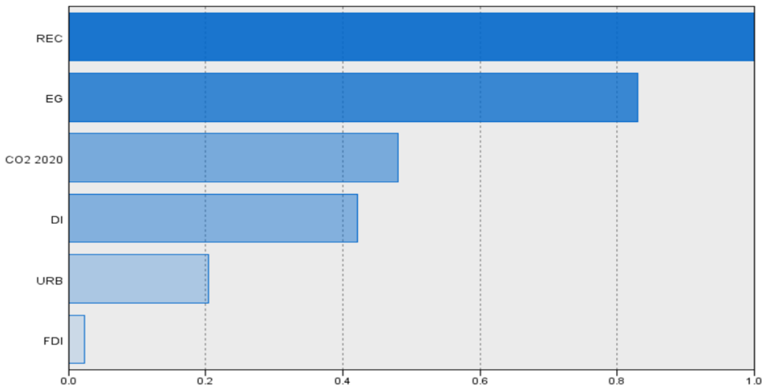
The graphical representation of variables contributing to the cluster is shown in Figure 3. The clusters are denoting in blue and yellow colour. Each dot is representing one country. The distinct dots represent good clustering except FDI in which blue and green dots seemingly overlapping each other.

Fig.3: Scatter of Cluster



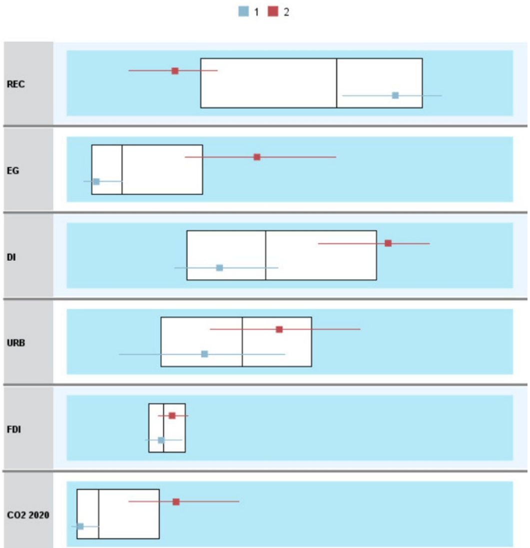
The role of variable in clustering can be understood by observing Figure 4 which shows the predictive importance of every variable is obtained on a scale of 0-1. The variable REC and EG is considered to be important variables followed by DI and URB. FDI emerges as the least important variable.

Fig 4: Predictive importance



The box-whisker plot corresponding to every variable is shown in Figure 5. The higher the distance among the red and blue whisker denoting the value at which variable is lying in the different cluster. CO2 variable is also being evaluated and it has been found that cluster 1 is having carbon emission at the lower side where cluster 2 is having more carbon emissions. Value of REC is on lower side in cluster 2 and hence establishing a relationship between carbon emissions and the renewable energy. Higher the level of renewable energy consumption, lower would be the carbon emissions.

Fig 5: Box and whisker plot of variables in clusters



The final cluster center is shown in Table 3 given below. Here all the variables are looking at distinct value in each cluster except FDI. FDI appears to occupy the same level in both the clusters (see Figure 5).

Table 3: Final Cluster Centers

Centroids											
		URB		FDI		REC		DI		EG	
		Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Cluster	1	37.71	16.77	4.68	8.5430	71.2	14.12	3.53	1.31	2976.73	1670.42
	2	52.01	19.19	3.56	2.8429	29.4	20.43	5.59	1.91	12110.8	6607.17
	Combined	43.1	18.85	4.26	6.9110	55.3	26.40	4.31	1.84	6456.38	6150.4

Clustering is effective if there is a significant difference in the values of the variable among the clusters. For understanding that some disruptive statistics like ANOVA is performed. It is found that except FDI all variables are significantly different in the cluster. Since clustering is performed to maximize the distance between the clusters, the ANOVA results can be used as descriptive statistics, and these have been shown in Table 4

Table 4: ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
CO2 2020	Between Groups	59.267	1	59.267	20.509	.000
	Within Groups	115.590	40	2.890		
	Total	174.856	41			
URB	Between Groups	2025.289	1	2025.289	6.452	.015

	Within Groups	12556.365	40	313.909		
	Total	14581.653	41			
FDI	Between Groups	12.418	1	12.418	.255	.616
	Within Groups	1945.821	40	48.646		
	Total	1958.239	41			
REC	Between Groups	17330.687	1	17330.687	61.613	.000
	Within Groups	11251.241	40	281.281		
	Total	28581.928	41			
DI	Between Groups	41.859	1	41.859	17.095	.000
	Within Groups	97.946	40	2.449		
	Total	139.805	41			
EG	Between Groups	826368638.352	1	826368638.352	45.619	.000
	Within Groups	724578643.553	40	18114466.089		
	Total	1550947281.905	41			

The mean and standard deviation of each variable is. shown in Table 3. To analyse the significant difference among the values of the variable in clusters, further t-test is performed.

The difference in t test along with the Levene's test is shown in Table 5 of independent t-test. Levene's test shows whether equal variability exists or not. The null hypothesis is that there is not a significant difference in the variances. Rejection of H_0 ($p<0.05$) implies that there exists a significant difference in the variance. Based on assumption of variances, t statistics is computed to analyse whether the level of variability is significantly different in both the clusters or not. The null hypothesis of t -test states that there is not a significant difference in the average level of mean between the clusters and rejection ($p<0.05$) implies that there exists a significant difference and hence it also implies that the variable is playing a role in clustering. The results also show that FDI is not lying at the different level in the cluster and hence plays different role in clustering.

Table 5: Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
CO2	EVA	15.924	.000	-4.5	40	.000	-2.44	.5401	-3.53	-1.35
	EVNA			-3.73	17.33	.002	-2.44	.6552	-3.82	-1.06
URB	EVA	.158	.693	-2.54	40	.015	-14.29	5.6296	-25.6	-2.92
	EVNA			-2.45	28.6	.020	-14.29	5.8170	-26.2	-2.39
FDI	EVA	3.457	.070	.505	40	.616	1.11	2.2162	-3.35	5.598
	EVNA			.615	33.02	.543	1.11	1.8199	-2.58	4.822

REC	EVA	.387	.537	7.84	40	.000	41.82	5.3290	31.05	52.60
	EVNA			7.19	23.88	.000	41.82	5.8108	29.83	53.82
DI	EVA	2.234	.143	-4.13	40	.000	-2.055	.49721	-3.06	-1.05
	EVNA			-3.78	23.7	.001	-2.055	.54320	-3.17	-.933
EG	EVA	20.605	.000	-6.75	40	.000	-9134.08	1352.355	-11867.2	-6400
	EVNA			-5.42	16.18	.000	-9134.08	1683.965	-12700.5	-5567

EVA: equal variance assumed; EVNA: equal variance not assumed

This leads to another important research question, if considering CO_2 as the variable which needs to be explained based on rest of the identified variables then what is the importance of variables in explaining the emission of CO_2. For this reason, feed forward multilayer neural network is devised. The relationship between the variables may be very complicated and hence can't be determined linearly. Multilayer perceptron neural network is formulated using the gradient descent algorithm. The type of training is batch, and the initial learning rate is set as 0.4 such that it will not skip too many minima points. To make it robust the training and testing is kept as 60% an 40% respectively. The model summary is given in Table 6 below:

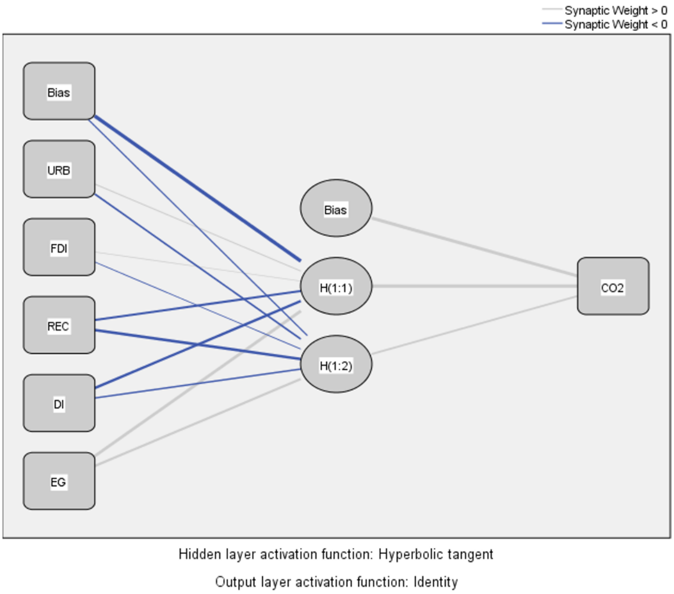
Table 6: Network Information

Input Layer	Covariates	1	URB
		2	FDI
		3	REC
		4	DI
		5	EG
	Number of Units[a]		5
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1[a]		2
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	CO2
	Number of Units		1
	Rescaling Method for Scale Dependents		Standardized
	Activation Function		Identity
	Error Function		Sum of Squares

a. Excluding the bias unit

The number of hidden layers in the neural network is found to be one and the number of units in the hidden layer are two. This implies neural network can deal with the complexity of the system and the deep learning is not required. The relationship between the variable and the synaptic weight is shown in Figure 6

Fig 6: Feed forward multilayer perceptron



The average importance score produced by the neural network on the basis of normalized importance is shown in Table 7 below.

Table 7: Relative Importance of Independent Variable

	Imp	N Imp	Imp	N Imp	Imp	N Imp	Imp	N Imp	Imp	N Imp	Average Imp	Rank
URB	.171	47.4	.150	39.3	.153	29.0	.084	21.1	.041	7.2	28.8%	4
FDI	.084	23.2	.035	9.1	.162	30.7	.224	56.2	.092	16.0	27.0%	5
REC	.272	75.2	.256	67.0	.528	100.0	.190	47.5	.124	21.7	62.3%	2
DI	.112	31.1	.176	46.1	.137	25.9	.103	25.7	.169	29.5	31.7%	3
EG	.361	100.0	.382	100.0	.019	3.7%	.399	100.0	.573	100.0	80.7%	1

Imp: Importance, N Imp: Normalize Importance in Percentage

The above table clearly shows that economic growth (proxied by GDPPC) turns out to be the most important variable affecting CO_2 emissions as it occupies rank 1 because of computed average importance score. REC emerges as the second important variable (with rank 2), DI with rank 3 is third in importance in explaining the environmental quality (proxied by CO_2 emissions). URB occupies the penultimate rank. FDI holds least importance in this entire exercise.

4. Conclusions and Policy Implications with Future Scope

What is the relative importance of the drivers of CO_2 emissions? In the present study we followed novel approach to assess the relative importance of the drivers of CO_2 emissions viz. economic growth (EG), renewable energy consumption (REC), urbanization (URB), democracy index (DI) and foreign direct investment (FDI) for their relation to the response variable viz. CO_2 emissions. We did the natural grouping of environmental quality (proxied by CO_2 emissions) experienced in different countries. The exploratory clustering is performed using hierarchical clustering which shows that the entire set of countries are classified into two groups. Later K mean clustering validates the hierarchical clustering result. The importance of variables in impacting environmental quality (CO_2) is carried out /gauged by neural networks. It has been found that FDI is least important element in understanding the nature of CO_2 emissions. The results match with the cluster analysis approach too in which FDI is found to be at insignificant difference level among both the clusters. Hence our results indicate that FDI's of the countries are inconclusive to determine the impact on CO_2 emissions. We documented that economic growth is the most important variable affecting CO_2 emissions as it occupies rank 1 based on computed average importance score. REC emerges as the second important variable (with rank 2), DI with rank 3 is third in importance in explaining the environmental quality (proxied by CO_2 emissions). URB occupies the penultimate rank.

Therefore, the clear-cut policy messages of our study is that emerging economies without bothering much about the external factors like FDI and its impact on CO_2 emissions should try to put their house in order by chalking out strategies that are well in their control. The study makes it amply clear that growth matters and clean / green growth will matter even more (as EG occupies Rank 1and REC gets Rank 2 in terms of average importance score). But the role of political system/ institutions cannot be ignored (as the next important variable in terms of average importance score is DI) and underestimated. To sum up, the point we want to emphasize here is that the emerging economies from Asia and SSA must focus on growth and in fact clean growth by promoting REC. This cannot happen without political vision and commitment as the political system too will matter as reflected by our previous study (Joshi et al., 2024). In fact, our previous study brings out that less democratic regimes can play an important role in ensuring clean environment via improving EQ. The plausible reason for this is that stringent environmental policies and regulatory mechanisms can be developed and implemented in less democratic countries more easily than in democratic regimes where myopic voters can oust the parties in next elections because of stringent /tough environmental protection measures. Given the penultimate rank of URB, there is a need for policy makers to enact policies for urban planning, management and governance and spread awareness about responsible consumption as well as production in urban areas to minimize carbon footprints and to accelerate progress on the sustainable development goals. Since there are two clusters that are emerging in our study, it will be interesting to study and compare the policies and processes of the countries falling in the same cluster. This constitutes the scope for future research. More research in this area is recommended.

References

Adams,S., Nsiah,C. (2019). Reducing carbon dioxide emissions; Does renewable energy matter? Science of the Total Environment. 693, 133298.https://doi.org/10.1016/j.scitotenv.2019.07.094

Ab-Rahim, R., Xin-Di, T. (2021). The Determinants of CO2 Emissions in ASEAN+3 Countries. Journal of Entrepreneurship and Business, 4(1), 38–49. https://doi.org/10.17687/jeb.v4i1.94

Adebayo, T.S., Kalmaz, D.B. (2021). Determinants of CO2 emissions: empirical evidence from Egypt. Environ Ecol Stat. 28, 239–262. https://doi.org/10.1007/s10651-020-00482-0

Adebayo, T.S., Akinsola, G.D.; Odugbesan,J.A., Olanrewaju,V.O. (2021).Determinants of Environmental Degradation in Thailand: Empirical Evidence from ARDL and Wavelet Coherence Approaches. 7(1): 181-196.https://doi.org/10.22059/poll.2020.309083.885

Ahmad, N., Du, L., Lu, J., Wang, J., H.-Z., Li, M.Z. (2017). Modelling the CO 2 emissions and economic growth in Croatia: Is there any environmental Kuznets curve? Energy. 123, 164–172.
Bai, X. McPhearson, T. Cleugh, H. Nagendra, H. Tong, X., Zhu, T. ,Zhu, Y.G.(2017).Linking Urbanization and the Environment: Conceptual and Empirical Advances. Annual Review of Environment and Resources.42(1),215-240.http://dx.doi.org/10.1146/annurev-environ-102016-061128

Balogh, J.M., Jámbor, A. (2017). Determinants of CO2 emission: a global evidence. In: International Journal of Energy Economics and Policy. 7 (5),217 - 226. http://hdl.handle.net/11159/1311

Bernauer, T., Koubi, V.(2009). Effects of political institutions on air quality. Ecol. Econ. 68 (5), 1355–1365. https://doi.org/10.1016/j.ecolecon.2008.09.003

Copeland, B.R, Taylor M.S. (1994). North-South Trade and the Environment. Q J Econ. 1994;109(3):755–87. https://doi.org/10.2307/2118421.

Dulal, H.B., Shah, K.U., Sapkota, C., Uma, G., Kandel, B.R., 2013. Renewable energy diffusion in Asia: can it happen without government support? Energy Pol. 59, 301e311.

Forgy, E. W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. Biometrics. 21 (3): 768–769. JSTOR 2528559.

Glavas, D. (2024). Determinants of CO2 Emissions: A Machine Learning Approach. Int Adv Econ Res 30, 215–217. https://doi.org/10.1007/s11294-024-09892-3

Han, J.,Meng, X., Zhou, X., Yi,, B ., Liu, M., Xiang, W.N.(2017).Along-term analysis of urbanization process, landscape change, and carbon sources and sinks: a case study in China's Yangtze River Delta region. J. Clean. Prod. 141,1040–50. https://doi.org/10.1016/j.jclepro.2016.09.177

The International Energy Agency (IEA)(2019). Global CO2 Emissions in 2019. https://www.iea.org/articles/global-co2-emissions-in-2019

The International Energy Agency (2024). GDP and carbon dioxide emissions are diverging, but not fast enough. https://energy.wisc.edu/news/gdp-and-carbon-dioxide-emissions-are-diverging-not-fast-enough.

Iwata, H., Okada, K., Samreth, S. (2011). Empirical study on the determinants of CO2 emissions: evidence from OECD countries. Applied Economics, 44(27), 3513–3519. https://doi.org/10.1080/00036846.2011.577023.

Joshi, S., Kansil, R. (2024). Does digital adoption improve environmental quality and facilitate sustainability? Evidence from emerging economies. Ref. Modul. Soc. Sci. https://doi.org/10.1016/B978-0-44-313776-1.00221-X.

Joshi, S., Gupta, S., Kaistha., C. (2024). Renewable energy consumption and CO2 emissions; do political systems matter? Arguments and evidence for emerging economies from Asia and SSA. Reference Module in Social Sciences. https://doi.org/10.1016/B978-0-44-313776-1.00285-3.

Ketchoua, G.S., Arogundade, S. & Mduduzi, B. Revaluating the Sustainable Development Thesis: exploring the moderating influence of Technological Innovation on the impact of Foreign Direct Investment (FDI) on Green Growth in the OECD Countries. Discov Sustain 5, 252 (2024). https://doi.org/10.1007/s43621-024-00433-w

Khan M, Rana AT, Ghardallou W. (2023).FDI and CO2 emissions in developing countries: the role of human capital. Nat Hazards. 117(1):1125–55. https://doi.org/10.1007/s11069-023-05949-4.

McKinsey Global Institute(MGI) (2018). Outperformers: High-Growth Emerging Economies and the Companies that Propel them. https://www.mckinsey.com/~/media/mckinsey/industries/public%20and%20social%20sector/our%20insights/outperformers%20high%20growth%20emerging%20economies%20and%20the%20companies%20that%20propel%20them/mgi-outperformers-full-report-sep-2018.pdf

Morales-Lage,R.; Bengochea-Morancho,A., Martinez-Zarzoso,I.(2016).The determinants of CO2 emissions: evidence from European countries. https://www.doctreballeco.uji.es/wpficheros/Morales_et_al_04_2016.pdf

Moriarty, P. Honnery, D. (2012). What is the global potential for renewable energy? Enery reviews.16(1).244-252.

Nielsen, F. (2016). Hierarchical Clustering. In: Introduction to HPC with MPI for Data Science. Undergraduate Topics in Computer Science. Springer, Cham. https://doi.org/10.1007/978-3-319-21903-5_8

Ojha,V.K., Abraham, A., Snášel., V. (2017). Metaheuristic design of feedforward neural networks: A review of two decades of research. Engineering Applications of Artificial Intelligence. 60:,97–116. https://doi.org/10.1016/j.engappai.2017.01.013

Peter J. Rousseeuw, Silhouettes(1987). A graphical aid to the interpretation and validation of cluster analysis, Journal of Computational and Applied Mathematics, Volume 20, Pages 53-65, ISSN 0377-0427, https://doi.org/10.1016/0377-0427(87)90125-7.

Qiwen ,X., Hailin ,W., Xinzhe, L., Xunzhang, P. (2021) Drivers of global and national CO2 emissions changes 2000–2017, Climate Policy, 21:5, 604-615, DOI: 10.1080/14693062.2020.1864267

Shah, S.A.A., Shah, S.Q.A., Tahir., M. (2022). Determinants of CO2 emissions: exploring the unexplored in low-income countries. Environ Science and Pollution Research. 29, 48276–48284. https://doi.org/10.1007/s11356-022-19319-3.

Sarkodie, S.A.; Adams, S. (2018). Renewable energy, nuclear energy, and environmental pollution: Accounting for political institutional quality in South Africa. Science of the Total Environment,643,1590-1601. https://doi.org/10.1016/j.scitotenv.2018.06.320.

Surroop, D., Raghoo, P. (2018). Renewable energy to improve energy situation in African island states. Renewable and Sustainable Energy Reviews. 88, 176–183. https://doi.org/10.1016/j.rser.2018.02.024.

Tukhtamurodov A, Sobirov Y, Toshalieva S, Ibrayimova D, Feruz M. (2024) Determinants of CO 2 emissions in the BRICS. A dynamic Panel ARDL approach. BIO Web Conf. 82:06002. https://doi.org/10.1051/bioconf/20248206002.

Wang., Q. (2014). Effects of urbanization on energy consumption in China. Energy Policy .65,332–9. https://doi.org/10.1016/j.enpol.2013.10.005

Wang, Q., Wu, S., Zeng, Y., Wu, B. (2016).Exploring the relationship between urbanization, energy consumption, and CO2 emissions in different provinces of China, 54(0),1563-1579. https://doi.org/10.1016/j.rser.2015.10.090

Wang, S., Fang, C., Guan, X, Pang, B., Ma, H., (2014). Urbanisation, energy consumption, and carbon dioxide emissions in China: a panel data analysis of China's provinces. Applied Energy. 136,738–49. https://doi.org/10.1016/j.apenergy.2014.09.059

Wang, Y., Chen, L., Kubota, J. (2016). The relationship between urbanization, energy use and carbon emissions: evidence from a panel of Association of Southeast Asian Nations (ASEAN) countries. Journal of Cleaner Production. 112(Part 2),1368–74. https://doi.org/10.1016/j.jclepro.2015.06.041

Xia, Q., Wang, H., Liu, X., & Pan, X. (2021). Drivers of global and national CO2 emissions changes 2000–2017. Climate Policy, 21(5), 604–615. https://doi.org/10.1080/14693062.2020.1864267

Zhang, N., Yu, K., Chen, Z. (2017). How does urbanization affect carbon dioxide emissions? A cross-country panel data analysis. Energy Policy. 107:678–687. https://doi.org/10.1016/j.enpol.2017.03.072

Zheng, H., Hu, J., Wang, S., & Wang, H. (2019). Examining the influencing factors of CO2 emissions at city level via panel quantile regression: evidence from 102 Chinese cities. Applied Economics, 51(35), 3906–3919. https://doi.org/10.1080/00036846.2019.1584659

Zmami, M., & Ben-Salha, O. (2020). An empirical analysis of the determinants of CO2 emissions in GCC countries. International Journal of Sustainable Development & World Ecology, 27(5), 469–480. https://doi.org/10.1080/13504509.2020.171550

Appendix

Table : Cluster Membership using K mean Clustering

Case Number	Countries	Cluster	Distance	Case Number	Countries	Cluster	Distance
1	Bangladesh	2	1767.26	22	Dem Rep of Congo	2	2629.313
2	Cambodia	2	374.456	23	Rep of Congo	2	806.999
3	China	1	959.651	24	Cote D Ivorie	2	1511.968
4	Fiji	2	6275.016	25	Equatorial Guines	1	4470.513
5	India	2	2885.042	26	Ethiopia	2	1344.46
6	Indonesia	1	5078.449	27	Gabon	1	3178.211
7	Lao PDR	2	3994.934	28	The Gambia	2	1533.356
8	Malaysia	1	10113.594	29	Guinea	2	1224.102
9	Myanmar	2	146.171	30	Guinea Bissau	2	1797.268
10	Nepal	2	173.282	31	Kenya	2	766.054
11	Papua New Guinea	2	305.122	32	Lesotho	2	1005.37
12	Philippines	2	5215.027	33	Liberia	2	2416.319
13	Sri Lanka	1	3966.722	34	Malawi	2	2239.208
14	Thailand	1	485.677	35	Mali	2	1572.101
15	Vietnam	2	4162.136	36	Mauritius	1	5480.62
16	Angola	2	1761.172	37	Mozambique	2	2507.323
17	Benin	2	296.737	38	Niger	2	2465.262
18	Botswana	1	346.636	39	Nigeria	2	1085.175

19	Burkina Faso	2	1587.126	40	Rwanda	2	1495.361
20	Burundi	2	2973.27	41	Senegal	2	362.055
21	Cameroon	2	87.777	42	Sierraleone	2	2083.131

Table : Cluster member in 2 stage clustering.

Countries	cluster number	Countries	cluster number
Dem Rep of Congo	1	Senegal	2
Burundi	1	Bangladesh	2
Niger	1	Lesotho	2
Malawi	1	Sri Lanka	2
Sierraleone	1	Philippines	2
Guinea Bissau	1	Fiji	2
Rwanda	1	India	2
Ethiopia	1	Gabon	2
Mozambique	1	Indonesia	2
Liberia	1	Vietnam	2
Burkina Faso	1	Botswana	2
Mali	1	Mauritius	2
The Gambia	1	Thailand	2
Guinea	1	Equatorial Guines	2
Cameroon	1	China	2
Kenya	1	Malaysia	2
Cote D Ivorie	1		
Rep of Congo	1		
Benin	1		
Nepal	1		
Nigeria	1		
Myanmar	1		
Papua New Guinea	1		
Angola	1		
Cambodia	1		
Lao PDR	1		