From Farm to Fork: Advanced Analysis, Visualization, and Imputation of CO2 Emissions in Food Production

Dhrsaj Kaushal, *Member, IEEE*
School of Engineering
Ajeenkya D Y Patil University
Pune, India
dhrsaj.kaushal@adypu.edu.in

***Abstract*—In the quest for sustainability, understanding and mitigating the carbon footprint of food products is paramount. This study aims to address this pressing need by conducting a comprehensive assessment of the carbon footprints of food products. We use advanced analytical techniques to rigorously analyze the complicated network of carbon emissions across the whole food supply chain, including the production and distribution phases. By precisely estimating greenhouse emissions, our study intends to provide a comprehensive knowledge of the carbon landscape, identifying important contributors and potential mitigation solutions. To aid comprehension and decision-making, advanced visualization approaches are used to transform complex data into intuitive representations, revealing spatial and temporal patterns of emission distribution and hotspots. Furthermore, in cases where data integrity is jeopardized due to missing or incomplete information, intensive imputation procedures are used to improve accuracy and completeness. Our research promotes scientific knowledge by combining these techniques and providing actionable insights for stakeholders throughout the food business spectrum. By providing decision-makers with a better understanding of carbon emissions and their drivers, this study helps ongoing efforts to build a more sustainable food system.**

***Index Terms*—Sustainability, carbon footprint, food products, analysis, visualization, prediction, model**

I. Introduction

The global food system stands at the intersection of two critical challenges facing humanity: ensuring food security for a growing population while mitigating the impacts of climate change [1]. With the urgency to reduce greenhouse gas emissions becoming increasingly apparent, the imperative to understand and mitigate the carbon footprint of food products has risen to the forefront of sustainability discourse. The food industry's substantial contribution to global greenhouse gas emissions, which come from the production, processing, transportation, and consumption phases of the food supply chain, emphasizes the severity of this issue even more.

To tackle this urgent requirement, a comprehensive and methodical evaluation of the carbon footprints linked to food items is necessary. Such an assessment necessitates the deployment of advanced analytical techniques capable of unraveling the complex network of carbon emissions across the entire food supply chain. By meticulously scrutinizing emissions from each stage of production and distribution, researchers can discern the primary contributors to carbon emissions and formulate effective mitigation strategies [2].

Furthermore, To guide policy interventions and decision-making processes aimed at promoting a more sustainable food system, precise measurement of greenhouse gas emissions is extremely important. In this regard, advanced visualization approaches play a pivotal role in translating intricate emission data into digestible representations, thus unveiling spatial and temporal patterns of emission distribution and identifying emission hotspots [3].

But evaluating carbon footprints is fraught with problems pertaining to data integrity, such as incomplete or missing data. Intensive imputation techniques are therefore necessary to improve the precision and comprehensiveness of emission estimations.

In light of these considerations, this study endeavors to undertake a comprehensive assessment of the carbon footprints of food products. Utilizing advanced analytical tools, we intend to thoroughly examine emissions along the food supply chain, from production to consumption. Notably, our predictive approach represents a substantial leap in this field by offering results in seconds, giving decision-makers immediate information to inform sustainable plans. Moreover, our model boasts high accuracy, validated through rigorous testing against empirical data, ensuring reliable and robust estimation of carbon footprints which allows us to work on projection models if we intend to in the future.

By providing a deeper understanding of carbon emissions and their underlying sources, this study hopes to help to ongoing efforts to construct a more resilient and sustainable food system in the face of climate change.

II. Literature Review

A. Reducing food's environmental impacts through producers and consumers.

Poore and Nemecek's thorough research emphasizes the importance of coordinated efforts between producers and consumers to address the environmental consequences of food production and consumption. The study provides actionable ideas for lowering emissions across the food supply chain. (Poore & Nemecek, 2018) [4].

B. Mitigating the greenhouse gas emissions embodied in food through realistic consumer choices.

This study focuses on the role of consumer behaviour in lowering the carbon footprint of food products, emphasizing the necessity of making informed dietary choices in reducing greenhouse gas emissions related to food consumption. (Hoolohan, McKinstry-West & Berners-Lee, 2013) [5].

C. Protein efficiency per unit energy and per unit greenhouse gas emissions: Potential contribution of diet choices to climate change mitigation.

By analyzing the environmental implications of dietary choices, including the transportation-related emissions of food, the study offers insights into the potential contribution of diet shifts to climate change mitigation efforts. The findings underscore the importance of considering the entire lifecycle of food products, including transportation, when assessing their environmental impacts and promoting sustainable dietary patterns.(Gonzalez, Frostell & Carlsson-Kanyama, 2011) [6].

III. Research Methodology

A. Objectives of the study

- Understanding which types of food have a more negative impact on the environment.
- To understand which stage of food production contributes more to greenhouse gas emissions.
- To compare the carbon footprint of plant-based and animal-based foods.

B. Research Design

The dataset utilized in this study was sourced from 'Our World in Data', a reputable platform that provides comprehensive datasets and visualizations on global trends and issues. 'Our World in Data' is known for its rigorous data collection methods and commitment to transparency, making it a trusted source for researchers and policymakers seeking reliable data on various socio-economic and environmental indicators.

C. Tools of data collection

The dataset, which comprises the 43 most common foods grown across the globe and their respective land, water usage, and carbon footprints., was accessed from https://ourworldindata.org/environmental-impacts-of-food

D. Data Interpretation techniques

- Statistical Analysis: calculations such as mean, median, standard deviation, regression analysis, etc.
- Data Preprocessing: preprocessing steps such as normalization, scaling, feature engineering, etc. were performed.
- Data Visualization: various plots and charts were used to visualize different aspects of the data.
- Machine Learning Algorithms: algorithms such as decision trees, random forest, AdaBoost, and K-nearest were included.

IV. Results and Observations

A. Total Emissions by Foods (Fig. 1.)

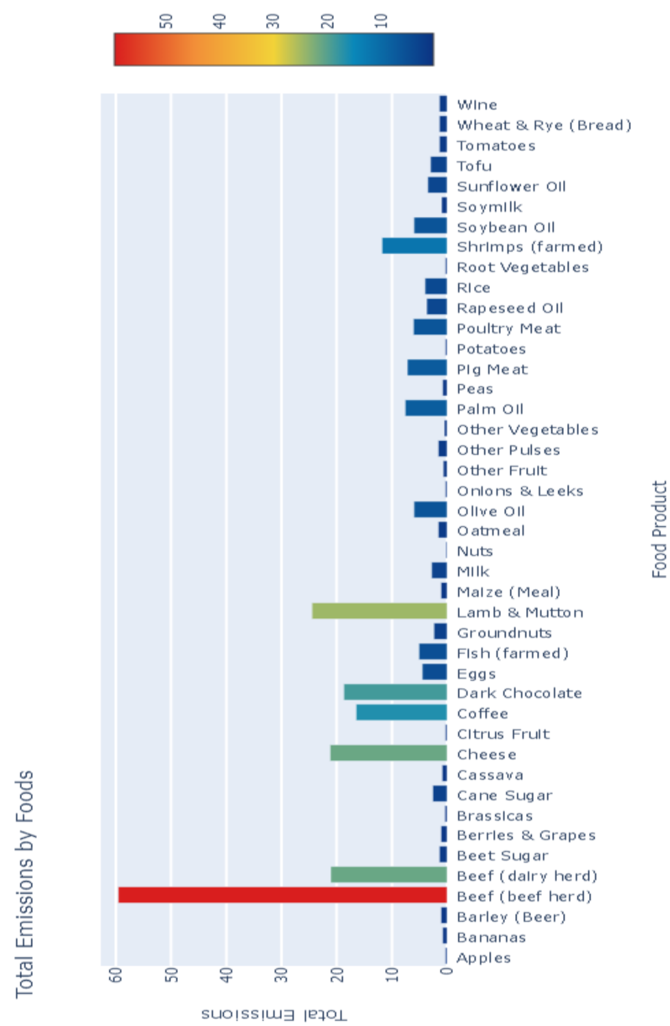

Fig. 1. Total Emissions by Foods.

1) Looks like Beef is responsible for most of the greenhouse gas emissions.
2) Plant-based foods are hardly visible in the graph indicating their significantly low carbon footprint.

B. Greenhouse gas emissions across the supply chain (Fig. 2.)

1) CO2 emissions from most plant-based Foods are much lower than most animal-based foods.

C. Land use per thousand calories (Fig. 3.)

1) CO2 emissions from most plant-based Foods are much lower than most animal-based foods.

D. Greenhouse emissions by Food category (Fig. 4.)

1) The greatest disparity appears in Farm emissions where Animal Products dominate with an astonishing 60.5 percent of emissions. Packaging and Processing emissions

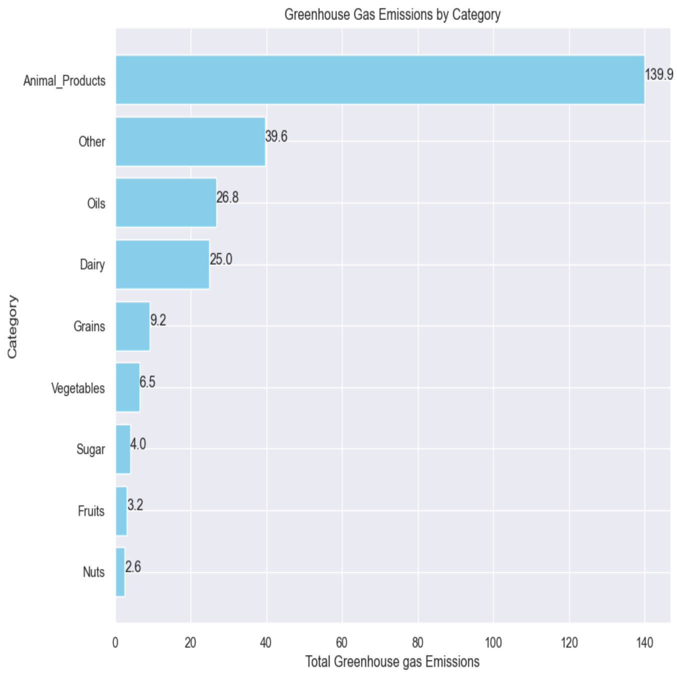

Fig. 2. Emissions per Food category.

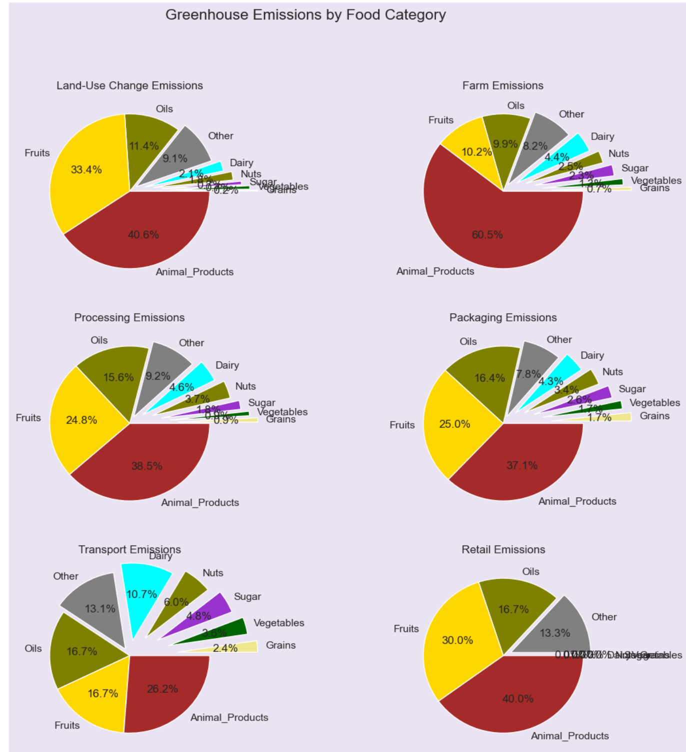

Fig. 4. Land Use per 1000 Calories.

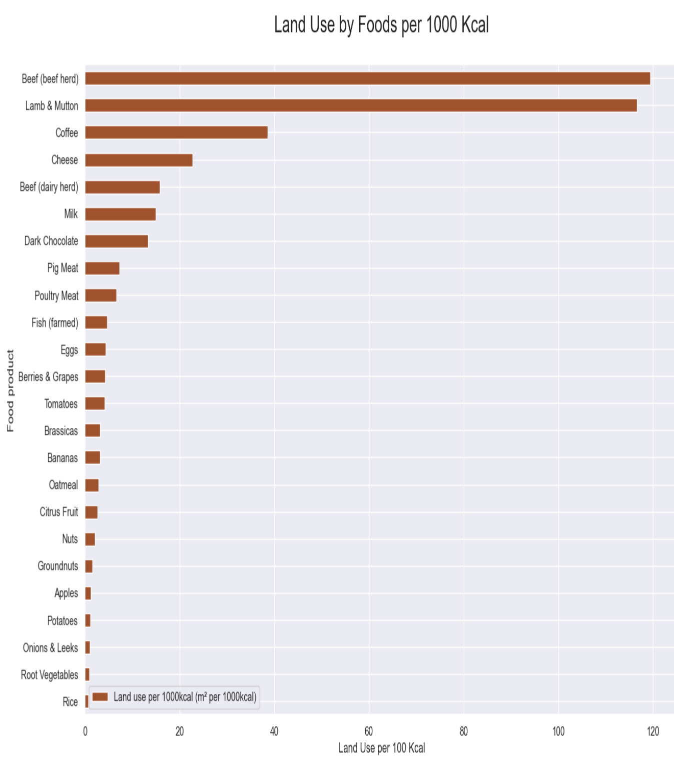

Fig. 3. Land Use per 1000 Calories.

appear to be relatively more balanced, while Transport emissions are by far the most equal across the board. Surprisingly, retail emissions appear to concentrate around three major food categories: Oils, Fruits, and Animal Products with a noticeable lack in most other categories.

E. Fresh water use by foods per Kg (Fig. 5.)

1) Cheese and Nuts need more water to produce the same amount(1 kg) of food than notable items like fish and beef.
2) Potatoes, root vegetables, coffee, and Onions are some of the foods that require very little water to produce.

F. Eutrophication emissions per Kg (Fig. 6.)

1) Animal-based foods are most responsible for eutrophication emission to produce 1kg while plant-based foods contribute very less.

G. Greenhouse gas emissions per 100g protein (Fig. 7.)

1) Dark chocolate has the most carbon footprint regarding nutritional values(per 100g protein), more than double that of the next item on the list.

H. Prediction (Fig. 8)

Here, the dataset was split into a training set and a testing set (90 percent train and 10 percent test) and scaling was done using StandardScaler. Various regression techniques were

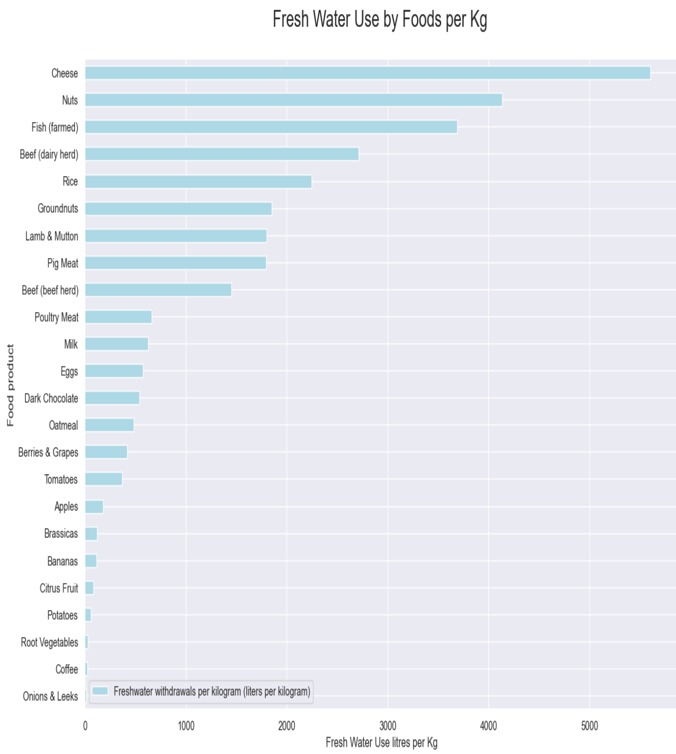

Fig. 5. Fresh Water Use per Kg.

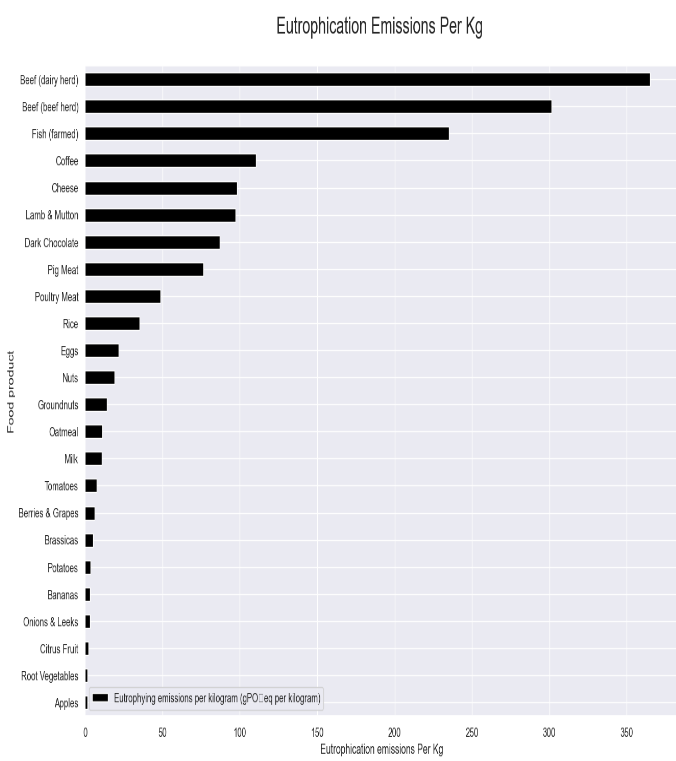

Fig. 6. Eutrophication Emissions per Kg.

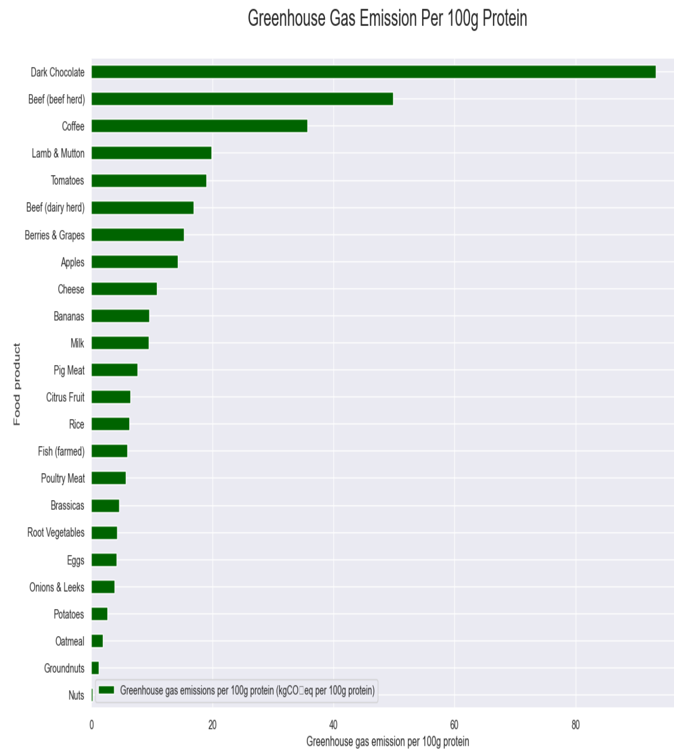

Fig. 7. Greenhouse Gas Emissions per 100g Protein.

applied on the training set to make predictions on the testing set.

1) Hence from the above comparison of the models in terms of R2 score we can see that AdaBoost Regression model outperformed all the other models by giving the highest R2 of 96.06 percent and RandomForest model comes in the second position by giving an R2 score of 90.41 percent.

I. Predictive Accuracy (Fig. 9)

1) The graph backs up our AdaBoost regressor accuracy by having a mean squared error of less than 3 percent.

V. Future Implications

The development and implementation of a model that identifies areas of resource wastage in food production and assesses the potential impact of resource-saving measures carry significant implications for the future of sustainable agriculture and environmental stewardship. Here, we delve into the potential future implications of such a model which will provide us with specific solutions to:

- Enhance Resource Efficiency
- Reduce Emissions
- Sustainable Land Management
- Conserve Water
- Economic Benefits
- Policy and Regulatory Implications

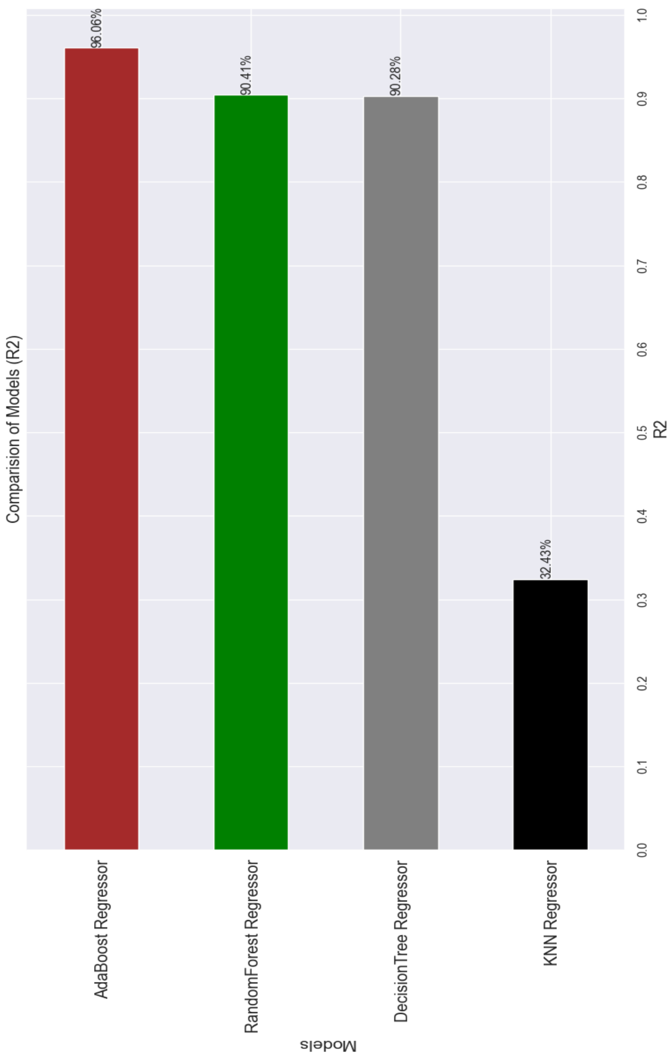

Fig. 8. Comparison of Models.

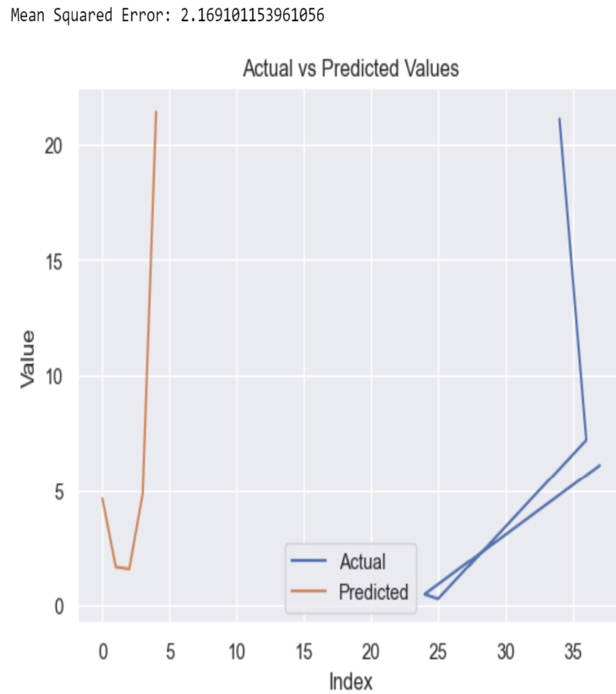

Fig. 9. Actual vs Predicted Values.

VI. Conclusion

This study has demonstrated the critical importance of understanding the carbon footprint of food products in the context of sustainability. Through a comprehensive assessment of carbon emissions across the food supply chain, we have gained valuable insights into the key contributors to greenhouse gas emissions and identified potential avenues for mitigation.

Our research has highlighted the significance of advanced analytical techniques and visualization approaches in elucidating complex emission data and informing decision-making processes. By leveraging these tools, we have been able to provide timely and accurate estimations of carbon footprints, facilitating the development of evidence-based sustainability strategies.

Moreover, our predictive model offers rapid results with high accuracy. This not only enhances the efficiency of carbon footprint assessments but also strengthens the reliability of our findings, thereby supporting more informed decision-making by stakeholders across the food sector.

Looking ahead, the findings of this study have important implications for policy development, business practices, and consumer behavior. Ultimately, by fostering collaboration and innovation within the food industry and beyond, we can pave the way for a more sustainable future where food production contributes to, rather than detracts from, environmental well-being. As we continue to strive towards this goal, the insights generated by this research will serve as a valuable resource for guiding efforts to build a more sustainable and equitable world for future generations.

Acknowledgment

I would like to express my sincere gratitude to Ajeenkya D Y Patil University for their support in this research project. I also want to express my gratitude to Prof. Khushbu Trivedi, our supervisor, for all of the help and advice during the research project. The project had some restrictions, but with our professor's guidance, we were able to get beyond them. I have learned a great deal from this research, both personally and professionally. Lastly, I want to express my sincere gratitude to Our World in Data for providing me with all the data I needed to do the study.

References

[1] Vermeulen, S. J., Campbell, B. M., Ingram, J. S. (2012). Climate change and food systems. Annual Review of Environment and Resources, 37, 195-222.

[2] Garnett, T. (2011). Where are the best opportunities for reducing greenhouse gas emissions in the food system (including the food chain), Food Policy, 36, S23-S32.

[3] Tilman, D., Clark, M. (2014). Global diets link environmental sustainability and human health. Nature, 515(7528), 518-522.

[4] Poore, J., Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. Science, 360(6392), 987-992.

[5] Hoolohan, C., Berners-Lee, M., McKinstry-West, J. (2013). Mitigating the greenhouse gas emissions embodied in food through realistic consumer choices. Energy Policy, 63, 1065-1074.

[6] Gonzalez, D., Frostell, B., Carlsson-Kanyama, A. (2011). Protein efficiency per unit energy and per unit greenhouse gas emissions: Potential contribution of diet choices to climate change mitigation. Food Policy, 36(5), 562-570.