Received: 22 April 2025 | Revised: 31 July 2025 | Accepted: 3 September 2025 | Published online: 26 September 2025

RESEARCH ARTICLE

Journal of Computational and Cognitive Engineering
2025, Vol. 00(00) 1-7
DOI: 10.47852/bonviewJCCE52025984

AI-Driven Prediction of Greenhouse Gas Emissions in Livestock Supply Chains: Towards a Data-Driven Model for Sustainable Agri-food Systems

Baha M. Mohsen[1,*] and Mohamad Mohsen[2]

[1] *Faculty of Business Management, Emirates Aviation University, United Arab Emirates*

[2] *College of Business, Eastern Michigan University, USA*

Abstract: This greenhouse gas (GHG) emissions mainly from enteric fermentation and manure management in livestock are mostly from copious methane (CH_4) and some nitrous oxide (N_2O). IPCC (International Panel on Climate Change) Tier 1 and Tier 2 emission prediction methods cannot provide details about where the emissions come from and are ineffective for different forms of agriculture. In our research, a machine learning model based on national agricultural data forecasts how much CH_4 and N_2O will be emitted from U.S. manure management. The performance of Random Forest Regression (RFR), Extreme Gradient Boosting (XGBoost) and Support Vector Regression (SVR) was tested by using RMSE, MAE and R^2 as performance metrics. XGBoost performed better than SVR since its predictive results were better than reaching $R^2 = 0.98$. The analysis of feature importance found that livestock type, methods of managing manure and population density are the main factors leading to emissions. The models resulted in information that communities in various locations could use to improve their sustainability. An adaptable decision-making procedure is proposed by the research to assist environmental planning in the agri-food sector and to ensure that intelligent agricultural platforms can better manage GHG emissions. More research is needed to improve the model by studying additional aspects from the supply chain, covering both its upstream and downstream operations, to obtain a complete analysis of environmental results. Future work should aim to incorporate additional stages of the livestock supply chain and adopt explainable AI techniques to improve transparency and support real-time decision-making.

Keywords: machine learning, greenhouse gas emissions, livestock supply chain, Random Forest Regression, XGBoost, sustainable agriculture, AI

1. Introduction

The basic food production network faces rising criticism due to the environmental destruction it causes, mainly through greenhouse gas (GHG) emissions. Livestock production systems generate extensive emissions that position them as key contributors among the diverse elements of this sector. According to the Food and Agriculture Organization (FAO), livestock activities generate 14.5% of all anthropogenic GHG emissions that mostly come from enteric fermentation and manure management releases [1, 2]. More recent studies such as that of Bilotto et al. [3] explain in more detail how the livestock industry has an important impact on global GHG emissions.

Animal agriculture faces rising environmental strains due to increasing product demand for meat and dairy, which is driven by population growth and changes in dietary preferences in developing nations combined with urbanization. Current trends require robust systems to monitor and minimize GHG emissions, as well as decision tools to aid supply chain sustainability decisions at different levels [4]. However, the IPCC (International Panel on Climate Change) Tier 1 and Tier 2 methodologies, combined with traditional estimation methods based on static emission factors, demonstrate limited ability for actual site-specific prediction of emissions because these approaches generalize across various production settings [5, 6]. Researchers now think that improved technologies may help measure GHG emissions more accurately in livestock farming. For example, McNicol et al. [7] suggest that the use of precision livestock technology can help control GHG emissions relating to beef. Bista et al. [8] further found that predicting GHG emissions in diversified semi-arid cropping systems is more accurate when site-specific data are used.

Currently, researchers see artificial intelligence (AI) and machine learning (ML) as major aids in improving environmental science models for farming. Complex, nonlinear relationships within different dimensions can be interpreted using these approaches and they also provide alternative choices to traditional fixed models that use data. Therefore, foreseeing growth in GHG emissions becomes more accurate and these tools allow operations to keep improving thanks to data learning [9, 10].

Adding AI to emission modeling systems helps achieve sustainability aims in both supply chain and logistics systems related to emission cuts and practical implementation. Predictive analytics helps organizations find where emissions are coming from, design the best resource plans, keep track of outcomes, and respond instantly to changes in environmental logistics [11, 12]. Furthermore, Ojadi et al. [13] point out that relying on big data analytics and AI enables companies to constantly check and predict which allows them to improve decision making and use less energy.

***Corresponding author:** Baha M. Mohsen, Faculty of Business Management, Emirates Aviation University, United Arab Emirates. Email: baha.mohsen@eau.ac.ae

© The Author(s) 2025. Published by BON VIEW PUBLISHING PTE. LTD. This is an open access article under the CC BY License (https://creativecommons.org/licenses/by/4.0/).

The practical implementation of ML technology for predicting livestock-generated GHG emissions exists only within restricted boundaries. Current applications dealing with agricultural systems mostly concentrate on crop systems while utilizing AI to optimize yield production and predict risks; however, animal agricultural emission sources are neglected. Moreover, modern ML approaches for practical decision support lack the integration of operational livestock data with manure management practice data from structured frameworks [14, 15]. Study teams have begun to discuss and investigate these gaps. For example, Symeon et al. [16] looked at different ways of dealing with manure, like anaerobic digestion and composting and stated how they may reduce GHG emissions in livestock farms.

To address the need, the author develops three ML models using supervision: Support Vector Regression, Random Forest, and Extreme Gradient Boosting (XGBoost). Data from national livestock records is used in these models to support manure management. The main point is to investigate how reliably these models predict CH_4 and N_2O emissions and whether they offer help in achieving sustainability in livestock supply chains.

2. Literature Review

Modern agricultural and environmental science fields demonstrate how livestock production creates GHG emissions. The livestock sector emits 14.5% of human-caused GHG emissions via CH_4 and N_2O released from digestive processes and proper waste management of domestic animals [1, 2]. The measurement of these emissions needs to be precise, as it facilitates better development of mitigation practices and helps livestock operations fulfill their climate obligations. The Tier 1 and Tier 2 emission assessment methods developed by the IPCC prove ineffective due to their dependency on set emission factors, as they do not account for the varied operational approaches across regions in livestock systems [5].

Researchers increasingly study AI and ML techniques to upgrade agricultural GHG emission modeling systems due to their capability to deliver improved precision, suitability, and prediction strength. These techniques have proven effective in handling complex nonlinear patterns and uniting diverse datasets between environmental zones and operational regions [10, 17]. Random Forest, Gradient Boosting, and Support Vector Regression have proven effective in determining agri-environmental predictions, such as crop yield estimation, irrigation scheduling, and nutrient management optimization [9, 18]. New research shows that ML can support the prediction of GHGs produced by farming. As a case in point, Toumi et al. [19] evaluated various ML regression models used for estimating GHG emissions, pointing out that the right model should be used, along with testing the importance of the features.

There is a lack of research in applying ML methods to monitor livestock GHG emissions, particularly regarding CH_4 and N_2O from enteric and manure systems. Most research focuses either on crop production systems or sustainability factors without making direct connections between empirical ML-based emission forecasting [20]. Caro et al. [21] introduce basic livestock emission patterns without implementing AI-based forecasting methods. Gollnow et al. [14] highlight the need for data-driven spatial methods to improve livestock emission inventories.

Developing supervised ML models through the integration of structured agricultural information about livestock populations, breeding methods, feedstuffs, and manure management allows for better emission forecasting outcomes [15, 22]. Predictive systems provide strategic capabilities for emissions-based logistics planning by helping to pinpoint emission centers in particular regions and optimize logistics resource deployment in supply chains. Thus, sustainable logistics principles advocating environmental stewardship must be integrated into operational and infrastructural decision-making [23]. In 2025, Toumi et al. [19] performed in-depth research on several ML regression algorithms and how they help understand and forecast GHG emissions for better sustainability.

AI models are being improved in supply chains through recent digital studies. Mohsen [24] explains that thanks to AI, the supply chain can forecast more accurately and is able to adapt more easily in its strategic decisions. Mohsen [25] explains that digital transformation supports flexible and strong logistics activities with the help of big data and intelligent automation. According to research, AI can make supply networks more efficient environmentally and responsive to customers [11, 12].

Even though these developments hold promise, there are still several problems. A gap in reliable emission readings from systems holds back further scrutiny and stakeholders are hesitant to use the models because they are still hard to explain. Model results adapted to specific production areas and systems are often needed for most of these models, as found by Kamilaris and Prenafeta-Boldú [10] and Odegard and van der Voet [26]. In order to meet these challenges, experts from both environmental data science and supply chain management should join efforts to develop effective solutions that apply to both areas. From new studies, it is evident that AI is needed to support more sustainable supply chains. The authors, Peeris and Baryannis [27], recommend guiding researchers in using AI transparently and responsibly to secure sustainable and equal circular supply chains. Meena et al. [28] state that AI makes it possible to automate activities in production and the supply chain for each of circularity's three principles.

To examine supervised ML models, the present research selects Random Forest, XGBoost, and Support Vector Regression, all based on well-structured livestock and manure management data. This study helps bridge a key understanding gap between AI, environmental effectiveness, and sustainability in animal markets.

3. Data Description

This research analyzes structured national statistics that connect livestock populations to GHG emissions released through manure management systems.

The livestock population data provides census records for animals categorized as dairy cattle, beef cattle, swine, sheep, and goats, along with data concerning manure management systems such as liquid/slurry, solid storage, and dry lot. Furthermore, the data follows national reporting guidelines. This segmentation facilitates the capture of operational variability in GHG emissions. These population numbers follow the same reporting standards applied by global inventory systems maintained by FAO [2] and EPA [6].

Yearly CH_4 and N_2O measurements expressed in CO_2-equivalent (CO_2e) are included in the emissions dataset together with standard emission factor and global warming potential (GWP) coefficient application. The available time frame of both datasets spans from 2012 to 2021, enabling the analysis of temporal variations and trend analysis. Data structures exist at the U.S. state levels through annual measurement periods, which create spatial details necessary for supervised learning to analyze both statewide patterns and annual change variations.

For the data to be reliable and high-quality, I had to perform many preprocessing activities during model development. In time series continuous variables, missing values were filled in by performing linear interpolation and forward-filling was implemented on the categorical proxy variables. Using the IQR approach, any points outside 1.5 times the IQR were singled out by the method, instead of just removing them as recommended by Khosravi et al. [29] for agri-environmental data.

Utilizing the feature engineering process, the models include livestock density metrics and kg CH_4 per head per year emission intensity. All numerical features were Z-score standardized to minimize the algorithm's learning bias about their size. To properly categorize the systems, they were assigned the one-hot encoding technique.

To conduct training and testing, the last set was divided using an 80:20 ratio. The data in this case was divided by using stratified sampling to make sure the numbers of livestock match the population. A review of the dataset formed the basis for the analysis and selection of models presented in the following method section.

4. Methodology

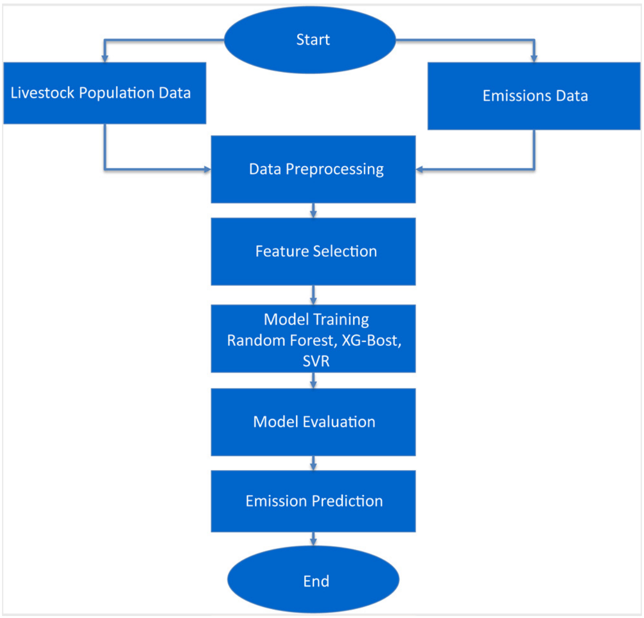
This work predicts the emissions of CH_4 and N_2O from livestock manure management by using ML algorithms. Farming research data was put together with durable ML algorithms to produce more focused, flexible, and understandable estimates of the environment's emissions. The main aim is to develop models that work well in a range of situations and times, as they support sensible decisions for the environment in agri-food systems.

Before assessing performance, the main steps of the research framework are data preprocessing, choosing and crafting useful features for engineering, choosing an algorithm, setting its hyperparameters, and validating. There are defined methods in place to assess whether models adhere to the guidelines in environmental modeling and ML applications [9].

Figure 1 shows the conceptual framework that directs every step of the ML pipeline this research uses. Start by cleaning your data, followed by studying how each algorithm operates, and finishing with appraising their explanations. The diagram outlines the breakdown of how emissions from livestock manure systems are estimated.

In this study, we considered both the type of data and what we knew about the domain. We began by doing a Pearson correlation analysis and removed features that were either repeated or had too much correlation with each other. The goal of feature selection was to choose the best variables by using both the feature rankings of a baseline Random Forest model and permutation importance, to prevent overfitting.

Figure 1
Conceptual framework for ML-based livestock GHG emission prediction



Environmental and agricultural modeling tasks benefit from the use of the following three regression algorithms, which were chosen for direct comparison:

1) RFR uses many decision trees in its learning which allows it to lower uncertainty in making predictions by combining their outputs. Because of robust overfitting prevention and the way it can model nonlinear relationships, the algorithm succeeds when working with complex environmental datasets [30].
2) It uses additive training as it trains boosting trees, giving the enhanced XGBoost its high scalability. Because of regularization and parallel computing, this model can handle the challenges of working with intricate agricultural data [31].
3) To address nonlinear patterns, Support Vector Regression first changes the input features into a high-dimensional form with kernel methods. This approach achieves good results in cases where data is hard to find or present unevenly [32].

Studies before ours demonstrated that these algorithms give the best outcomes in agro-environmental tasks, so we think they are appropriate for this research [33, 34].

The model's performance was measured using RMSE, MAE, and R^2 (coefficient of determination), all of which are explained by Equations (1)–(3) in the standard approach to regression metrics. Environmental ML experts say that these metrics can be used to gauge a model's accuracy in terms of bias, variance, and how much information the model passes on [35]. These ideas are shown in mathematical form as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_l\right)^2} \tag{1}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \widehat{y_l}\right| \tag{2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \widehat{y_l}\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \tag{3}$$

Where:

y_i: actual observed value
$\widehat{y_l}$: model-predicted value
$\bar{y}$: mean of the observed values
n: total number of observations

Ten-fold cross-validation helped confirm that the method would work well in general. Ten small groups were made from the training data, and the model was trained on nine of them to validate the remaining group. Grids and cross-validation were applied to find the ideal settings for parameters involved in the number of estimators, maximum tree depth for RFR and XGBoost, kernel type in SVR, and eps value for XGBoost. Since future application in operational systems would require these approaches to be efficient, this was considered an important factor in the research.

The optimal hyperparameters for each ML model were selected using grid search combined with ten-fold cross-validation. The final configurations used in the study are as follows:

1) Random Forest Regression (RFR):

Number of estimators = 300, maximum tree depth = 20, minimum samples split = 4, minimum samples leaf = 2, bootstrap = True.

2) Extreme Gradient Boosting (XGBoost):
Number of estimators = 250, maximum tree depth = 8, learning rate = 0.05, subsample = 0.8, colsample_bytree = 0.8, regularization alpha (L1) = 0.1, regularization lambda (L2) = 1.0.

3) Support Vector Regression (SVR):
Kernel = Radial Basis Function (RBF), C = 100, epsilon = 0.1, gamma = 'scale'.

These hyperparameter settings yielded the highest model performance based on the lowest RMSE and MAE and the highest R^2 values across validation folds.

All modeling procedures in this research were written and performed on Python version 3.10. The core ML models were generated with Scikit-learn and XGBoost libraries. This work was done by using Pandas, NumPy, and Matplotlib which are common tools and are recommended by current agricultural data science experts.

5. Results and Analysis

This section presents the prediction results of ML models designed to estimate CH_4 and N_2O emissions from livestock manure systems. The research compares three supervised regression algorithms – RFR, XGBoost, and SVR – to determine their effectiveness for emission prediction while analyzing feature importance and regional and management system variations. This analysis uses the Section 3 pre-processed dataset to support the model structure defined in Section 4.

5.1. Model performance comparison

Data evaluation was carried out using RMSE, MAE, and R^2, as defined in Equations (1)–(3). Table 1 shows how ensemble-based models delivered superior performance compared to SVR according to both error measurements and model predictiveness. XGBoost had the lowest RMSE and the highest R^2 while assessing both CH_4 and N_2O prediction targets, demonstrating superior prediction capability.

5.2. Feature importance analysis

The trained Random Forest and XGBoost models prompted the examination of feature importance rankings, which determined the most influential predictors of emissions from livestock operations. The primary influential features in the models were livestock type, the manure management system, and the livestock population density. The duration of animal manure management systems and climatic regions contributed to the main predictive factors. The results are consistent with previous research on environmental studies, demonstrating how stocking density and manure management methods affect emission profiles [1, 6].

Table 2 presents the top ten features ranked by importance based on the trained XGBoost model. These importance scores reflect the average gain in accuracy brought by each feature during the model's training process. The ranking confirms the significant influence of livestock type, manure management system, and population density on GHG emissions, aligning with prior empirical and theoretical expectations.

These insights can guide more effective interventions and policy development by identifying the most influential variables affecting livestock-based GHG emissions.

Table 1
Model performance comparison for CH_4 and N_2O prediction

Model	RMSE (CH_4)	MAE (CH_4)	R^2 (CH_4)	RMSE (N_2O)	MAE (N_2O)	R^2 (N_2O)
RFR	0.1362	0.0416	0.992	0.0442	0.0168	0.9886
XGBoost	0.1907	0.0442	0.9843	0.0335	0.0144	0.9935
SVR	0.408	0.0958	0.9281	0.0706	0.0520	0.9711

Table 2
Top 10 features ranked by importance (XGBoost model)

Rank	Feature	Importance score
1	Livestock type	0.223
2	Manure management system type	0.181
3	Livestock population density	0.153
4	Average temperature (regional)	0.102
5	Length of manure storage duration	0.089
6	Annual precipitation	0.076
7	CH_4 intensity (kg/head/year)	0.063
8	Type of manure storage (e.g., lagoon)	0.048
9	Emission factor (standardized)	0.038
10	Climate zone index	0.027

5.3. Emission forecasting by region and system type

Spatial variations were analyzed by aggregating predicted CH_4 and N_2O emissions at the U.S. state level, with data further categorized by manure management system types. The geographic analysis shows significant differences between states, with regions containing CAFOs and liquid/slurry-based manure systems producing the highest emissions. Area populations with high cattle numbers produced higher CH_4 emissions, while swine-intensive areas released higher N_2O from their anaerobic lagoons. Empirical analysis validates the need to direct specific mitigation strategies at particular livestock farming areas and management systems based on these predictions.

5.4. Model performance comparison

The model fit assessment included scatterplots that showed actual emission values against predicted values of the test dataset. Both Random Forest and XGBoost models showed effective agreement with observed values when evaluated through the clustering around the 45-degree reference line, as displayed in Figures 2 and 3. The linear distribution of SVR results appeared more widespread when prediction

Figure 2
Actual vs. predicted CH_4 emissions using Random Forest and XGBoost

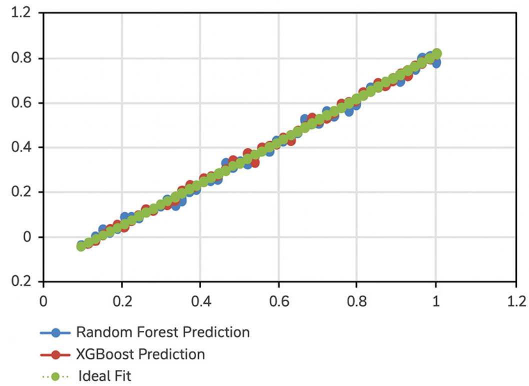
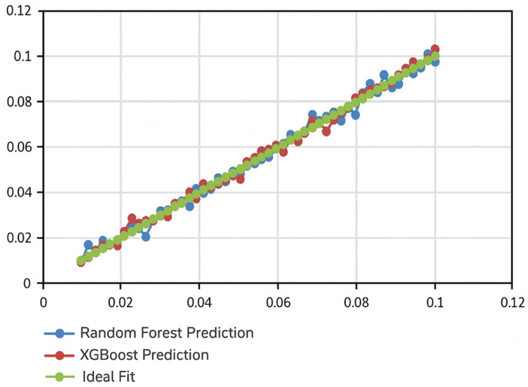


Figure 3
Actual vs. predicted N_2O emissions using Random Forest and XGBoost



values reached higher emission measurement points. The visual diagnostics support the numeric performance metrics, demonstrating that tree-based ensemble approaches have the best outcome in this application.

5.5. Interpretive insights

Around 23% of the overall contribution can be attributed to ensemble ML models which are strong at discovering complicated nonlinear patterns in GHG emissions from livestock. Emission reduction policies ought to target changes in infrastructure and work methods in crowded livestock production areas because these aspects have the greatest impact on manure management and livestock numbers. With localized forecasts, communities can create detailed plans for being sustainable by taking local projects into account.

Studies from earlier research agree that using data and decision-support tools helps with monitoring the environment for livestock farming [10, 14]. This approach hoped to improve learning and lead to better agri-food delivery network policies.

6. Discussion

Research demonstrates that using RFR and XGBoost as part of ML ensembles is very effective for forecasting GHG emissions caused by livestock waste management. Apart from their algorithms, the findings reveal that the models might support green methods for making decisions in the livestock supply chain.

6.1. Implications for sustainable livestock supply chains

Environmental accounting benefits from accurate CH_4 and N_2O emissions predictions since it becomes crucial for launching climate-smart agricultural practices. Stakeholders representing producers, regulators, and sustainability officers can direct investments towards emission reduction technologies by identifying their hotspots at the management system and state level [36, 37]. This method is consistent with global efforts by the FAO to develop sustainable livestock systems and is grounded in recent scientific recommendations by Gerber et al. [1] and Leip et al. [38] as its foundation.

6.2. Potential use in decision support systems

The ability of these ML models to explain and predict outcomes enhances their usefulness for decision support systems (DSS). Once these models are integrated into digital farm management systems, producers can assess the amount of greenhouse gases resulting from different farming choices right away. Moreover, these systems help regional planners and policymakers develop programs to reduce emissions by analyzing predictions. This neatly fits with new agriculture projects that encourage digital tools to measure environmental performance [10].

6.3. Comparison with previous studies

This research makes two important contributions to prior work in the field. Whereas previous studies by Gollnow et al. [14] included detailed inventorying of emissions, our method takes a different approach that maximizes accuracy in predictions rather than reporting each emission inventory. This second benefit is unlike other IPCC [6] methods, ML makes it simple to deal with changing data, allowing alignment with the call for more flexible approaches in agri-environmental science.

6.4. Limitations and future extensions

To build a more complete picture of environmental impact, future research should extend the modeling framework to encompass additional emissions sources beyond manure management. This includes upstream factors such as feed production (e.g., fertilizer application, deforestation), energy use on farms (e.g., ventilation, lighting), and downstream logistics and transportation. Incorporating these components would support a comprehensive life-cycle assessment (LCA) approach for livestock supply chains.

In addition, model transparency and practical utility could be significantly improved through the integration of Explainable AI (XAI) tools such as SHAP (SHapley Additive exPlanations) values. These methods help visualize and quantify the contribution of each input feature to individual predictions, making the results more interpretable for farmers, environmental regulators, and system designers.

Finally, to further validate the model's robustness and real-world applicability, future work should benchmark it against other ML architectures, including artificial neural networks (ANNs), recurrent neural networks (RNNs), and time-series models. Complementing this with validation using empirical field data from operational farms would strengthen confidence in the model's generalizability and encourage its integration into digital agriculture platforms.

In addition to the limitations mentioned, future work could significantly expand the scope of the current model by integrating additional sources of emissions across the livestock supply chain. These include emissions from upstream processes such as feed production (e.g., fertilizer application, land use change), on-farm energy use (e.g., for heating, cooling, or mechanized equipment), and downstream logistics and transportation activities. Incorporating these dimensions would allow for a more holistic LCA of environmental impacts.

Moreover, the predictive framework could be enhanced by integrating real-time monitoring technologies such as IoT-based sensors for measuring manure composition, temperature, and gas fluxes. These tools can provide dynamic, high-frequency data to improve the accuracy and responsiveness of the model. Likewise, employing high-resolution spatial data (e.g., at county or farm plot level) using GIS and remote sensing can refine spatial granularity, enabling localized policy recommendations and interventions.

Advancing the model in these directions would support a next-generation environmental DSS capable of proactive emission management in diverse livestock production systems.

As these issues are addressed, the predictive ML framework suggested here could turn into a complete environmental monitoring system supporting climate mitigation not only in the livestock supply chain but also across other sectors.

7. Conclusion

This work introduces a model that predicts both CH_4 and N_2O emissions from handling livestock manure with 10 years of U.S. state data. The two models with the greatest predictive accuracy are ensemble techniques RFR and XGBoost, recording R^2 values of more than 0.98. These models helped reveal that the main things affecting emissions are the type of livestock, the management system, and population density.

This study's most significant achievement is to add interpretable ML models to agri-environmental sustainability which helps with regional research and decision support. Through creating a flexible and highly accurate prediction method, the research fixes problems with old approaches that are based either on fixed emission data or wide assessments.

Policies and practices can be shaped using the model outputs by finding places with high emissions and by assessing the carbon effects of changing strategies. Adding the predictive framework to farm management software and environmental DSS platforms supports full disclosure, clear provenance, and a positive impact on climate change.

The scope of the model could be widened in future to study additional elements like supply chain feed delivery and effects of freight transport on climate change. Consequently, this work supports the effort to use AI to promote environmental sustainability in agriculture.

Extending this framework with life-cycle components such as feed production and logistics, and integrating tools like SHAP values, can further increase the model's impact and applicability in real-world smart agriculture systems.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in USDA at https://agdatacommons.nal.usda.gov/search?itemTypes=3&categories=30761,30746,30689,30821,33323.

Author Contribution Statement

Baha M. Mohsen: Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration. **Mohamad Mohsen:** Methodology, Validation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization.

References

[1] Gerber, P. J., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., ..., & Tempio, G. (2013). *Tackling climate change through livestock: A global assessment of emissions and mitigation opportunities*. Italy: Food and Agriculture Organization of the United Nations.

[2] Food and Agriculture Organization of the United Nations. (2021). https://www.fao.org/3/cb8685en/cb8685en.pdf

[3] Bilotto, F., Christie-Whitehead, K. M., Malcolm, B., Barnes, N., Cullen, B., Ayre, M., & Harrison, M. T. (2025). Costs of transitioning the livestock sector to net-zero emissions under future climates. *Nature Communications*, *16*(1), 3810. https://doi.org/10.1038/s41467-025-59203-5

[4] Springmann, M., Clark, M., Mason-D'Croz, D., Wiebe, K., Bodirsky, B. L., Lassaletta, L., ..., & Willett, W. (2018). Options for keeping the food system within environmental limits. *Nature*, *562*(7728), 519–525. https://doi.org/10.1038/s41586-018-0594-0

[5] Milne, A. E., Glendining, M. J., Bellamy, P., Misselbrook, T., Gilhespy, S., Rivas Casado, M., ..., & Whitmore, A. P. (2014). Analysis of uncertainties in the estimates of nitrous oxide and methane emissions in the UK's greenhouse gas inventory for agriculture. *Atmospheric Environment*, *82*, 94–105. https://doi.org/10.1016/j.atmosenv.2013.10.012

[6] United States Environmental Protection Agency. (2022). Inventory of U.S. greenhouse gas emissions and sinks. https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks

[7] McNicol, L. C., Bowen, J. M., Ferguson, H. J., Bell, J., Dewhurst, R. J., & Duthie, C.-A. (2024). Adoption of precision livestock farming technologies has the potential to mitigate greenhouse gas emissions from beef production. *Frontiers in Sustainable Food Systems*, *8*, 1414858. https://doi.org/10.3389/fsufs.2024.1414858

[8] Bista, P., Shakya, S. R., Acharya, P., & Ghimire, R. (2025). Evaluating machine learning models for greenhouse gas emissions prediction in diversified semi-arid cropping systems. *Soil Science Society of America Journal*, *89*(2), e70057. https://doi.org/10.1002/saj2.70057

[9] Pincheira, M., Vecchio, M., Giaffreda, R., & Kanhere, S. S. (2021). Cost-effective IoT devices as trustworthy data sources for a blockchain-based water management system in precision agriculture. *Computers and Electronics in Agriculture*, *180*, 105889. https://doi.org/10.1016/j.compag.2020.105889

[10] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, *147*, 70–90. https://doi.org/10.1016/j.compag.2018.02.016

[11] Dubey, R., Gunasekaran, A., Childe, S. J., Blome, C., & Papadopoulos, T. (2019). Big data and predictive analytics and manufacturing performance: Integrating institutional theory, resource-based view and big data culture. *British Journal of Management*, *30*(2), 341–361. https://doi.org/10.1111/1467-8551.12355

[12] Wamba, S. F., Dubey, R., Gunasekaran, A., & Akter, S. (2020). The performance effects of big data analytics and supply chain ambidexterity: The moderating effect of environmental dynamism. *International Journal of Production Economics*, *222*, 107498. https://doi.org/10.1016/j.ijpe.2019.09.019

[13] Ojadi, J. O., Odionu, C. S., Onukwulu, E. C., & Owulade, O. A. (2024). Big data analytics and AI for optimizing supply chain sustainability and reducing greenhouse gas emissions in logistics and transportation. *International Journal of Multidisciplinary Research and Growth Evaluation*, *5*(1), 1536–1548. https://doi.org/10.54660/.IJMRGE.2024.5.1.1536-1548

[14] Gollnow, S., Pieper, M., & Leip, A. (2022). A spatially explicit approach to quantify on-farm GHG emissions in European livestock systems. *Nature Food*, *3*(8), 677–686. https://doi.org/10.1038/s43016-022-00576-4

[15] Zheng, B., Cheng, J., Geng, G., Wang, X., Li, M., Shi, Q., Qi, J., Lei, Y., Zhang, Q., & He, K. (2021). Mapping anthropogenic emissions in China at 1 km spatial resolution and its application in air quality modeling. *Science Bulletin*, *66*(6), 612–620. https://doi.org/10.1016/j.scib.2020.12.008

[16] Symeon, G. K., Akamati, K., Dotas, V., Karatosidi, D., Bizelis, I., & Laliotis, G. P. (2025). Manure management as a potential

mitigation tool to eliminate greenhouse gas emissions in livestock systems. *Sustainability*, *17*(2), 586. https://doi.org/10.3390/su17020586

[17] Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big data in smart farming – A review. *Agricultural Systems*, *153*, 69–80. https://doi.org/10.1016/j.agsy.2017.01.023

[18] Zhou, J., Zhou, J., Ye, H., Ali, M. L., Chen, P., & Nguyen, H. T. (2021). Yield estimation of soybean breeding lines under drought stress using UAV-based imagery and convolutional neural networks. *Biosystems Engineering*, *204*, 90–103. https://doi.org/10.1016/j.biosystemseng.2021.01.017

[19] Toumi, S., Aljadani, A., Toumi, H., Ammouri, B., & Dhiabi, M. (2025). AI for climate change: Unveiling pathways to sustainable development through greenhouse gas emission predictions. *Eurasian Economic Review*. Advance online publication. https://doi.org/10.1007/s40822-024-00295-7

[20] Raj, N., Perumal, S., Singla, S., Sharma, G. K., Qamar, S., & Chakkaravarthy, A. P. (2022). Computer aided agriculture development for crop disease detection by segmentation and classification using deep learning architectures. *Computers and Electrical Engineering*, *103*, 108357. https://doi.org/10.1016/j.compeleceng.2022.108357

[21] Caro, D., Davis, S. J., Bastianoni, S., & Caldeira, K. (2014). Global and regional trends in greenhouse gas emissions from livestock. Climatic Change, *126*(1–2), 203–216. https://doi.org/10.1007/s10584-014-1197-x

[22] Rose, D. C., & Chilvers, J. (2018). Agriculture 4.0: Broadening responsible innovation in an era of smart farming. *Frontiers in Sustainable Food Systems*, *2*, 87. https://doi.org/10.3389/fsufs.2018.00087

[23] Mohsen, B. M. (2022). Principles of sustainable logistics. In S. Jerold Samuel Chelladurai, S. Mayilswamy, S. Gnanasekaran, & R. Thirumalaisamy (Eds.), *Logistics engineering* (pp. 1–12). IntechOpen. https://doi.org/10.5772/intechopen.103018

[24] Mohsen, B. M. (2023). Impact of artificial intelligence on supply chain management performance. *Journal of Service Science and Management*, *16*(1), 44–58. https://doi.org/10.4236/jssm.2023.161004

[25] Mohsen, B. M. (2023). Developments of digital technologies related to supply chain management. *Procedia Computer Science*, *220*, 788–795. https://doi.org/10.1016/j.procs.2023.03.105

[26] Odegard, I. Y. R., & van der Voet, E. (2014). The future of food—Scenarios and the effect on natural resource use in agriculture in 2050. *Ecological Economics*, *97*, 51–59. https://doi.org/10.1016/j.ecolecon.2013.10.005

[27] Peeris, M. P., & Baryannis, G. (2025). Artificial intelligence for supply chain sustainability: A research roadmap. In *16th International Conference on Software, Knowledge, Information Management and Applications*, 1–6.

[28] Meena, R., Sahoo, S., Malik, A., Kumar, S., & Nguyen, M. (2025). Artificial intelligence and circular supply chains: Framework for applications and deployment from the triple bottom line model perspective. *Annals of Operations Research*. Advance online publication. https://doi.org/10.1007/s10479-025-06510-1

[29] Khosravi, A., Nahavandi, S., Creighton, D., & Atiya, A. F. (2011). Lower upper bound estimation method for construction of neural network-based prediction intervals. *IEEE Transactions on Neural Networks*, *22*(3), 337–346. https://doi.org/10.1109/TNN.2010.2096824

[30] Breiman, L. (2001). Random forests. Machine Learning, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

[31] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

[32] Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, *14*(3), 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

[33] Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, *18*(8), 2674. https://doi.org/10.3390/s18082674

[34] Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., ..., & Kim, S.-H. (2016). Random forests for global and regional crop yield predictions. *PLoS ONE*, *11*(6), e0156571. https://doi.org/10.1371/journal.pone.0156571

[35] Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*, 79–82. https://doi.org/10.3354/cr030079

[36] Bocean, C. G. (2025). The role of organic farming in reducing greenhouse gas emissions from agriculture in the European Union. *Agronomy*, *15*(1), 198. https://doi.org/10.3390/agronomy15010198

[37] Olesen, J. E., Rees, R. M., Recous, S., Bleken, M. A., Abalos, D., Ahuja, I., ..., & Topp, C. F. E. (2023). Challenges of accounting nitrous oxide emissions from agricultural crop residues. *Global Change Biology*, *29*(24), 6846–6855. https://doi.org/10.1111/gcb.16962

[38] Leip, A., Weiss, F., Lesschen, J. P., & Westhoek, H. (2014). The nitrogen footprint of food products in the European Union. *The Journal of Agricultural Science*, *152*(S1), 20–33. https://doi.org/10.1017/S0021859613000786

How to Cite: Mohsen, B. M., & Mohsen, M. (2025). AI-Driven Prediction of Greenhouse Gas Emissions in Livestock Supply Chains: Towards a Data-Driven Model for Sustainable Agri-food Systems. *Journal of Computational and Cognitive Engineering*. https://doi.org/10.47852/bonviewJCCE52025984