Modelling methane emissions from rice paddies using machine learning

Abira Sengupta, Fathima Nuzla Ismail
University of Otago, New Zealand, State University of New York at Buffalo, USA
abira.sengupta@otago.ac.nz, fathima.nuzla.ismail@gmail.com

Aim

One significant contributor to global warming is the natural release of methane (CH_4) from rice paddies. This approach uses explainable AI (XAI) to predict CH4 emissions from rice paddies in West Bengal, India, using various features. This study highlights the importance of hyperparameter adjustment in enhancing model accuracy and identifying significant features.

Dataset

This study uses Copernicus Atmosphere Monitoring Service (CAMP) global greenhouse gas reanalysis (EGG4)[a] with $0.75^\circ \times 0.75^\circ$ spatial resolution.

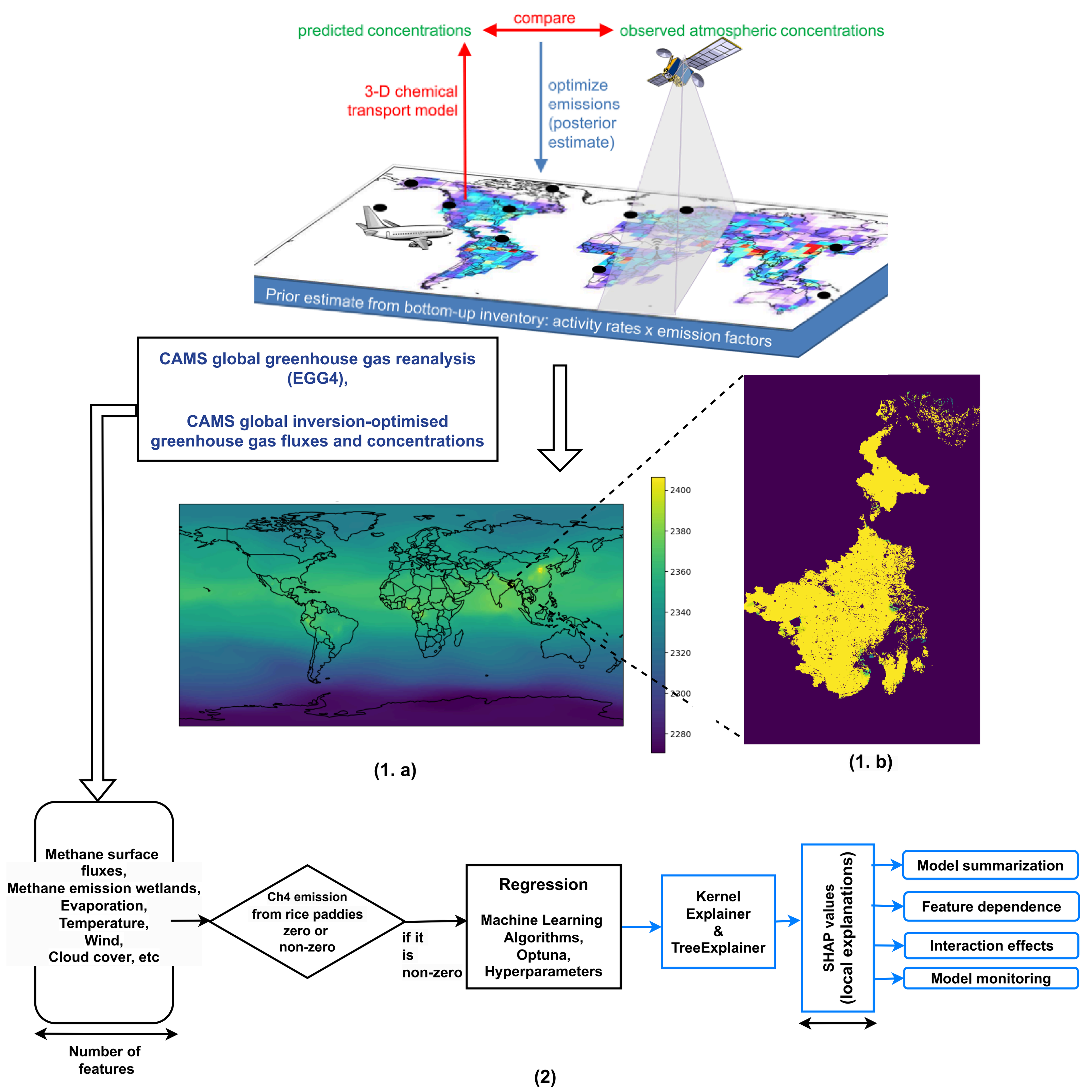

Figure: (1. a) represents worldwide CH_4 fluxes. (1. b) represents CH_4 emissions from rice paddies in West Bengal, India. (2) represents *KernelExplainer* and *TreeExplainer* enable a wide variety to understand the global model

[a] https://ads.atmosphere.copernicus.eu/datasets/cams-global-ghg-reanalysis-egg4?tab=documentation

Methane emission

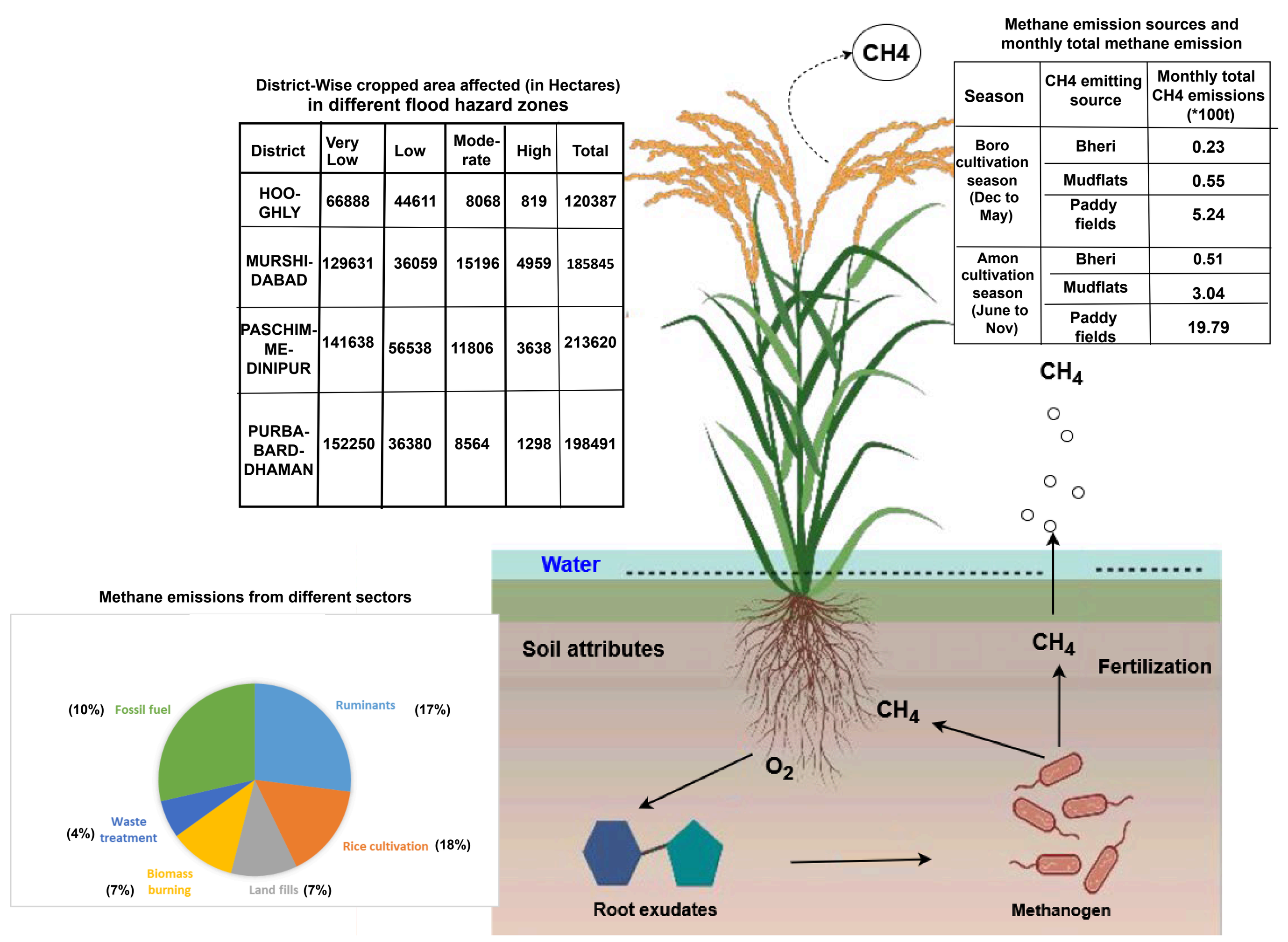

Figure: Factors affecting methane emissions from rice paddies[b] [2].

Methane emissions calculation [1]

$$\text{Methane emission (Gg yr}^{-1}) = \sum_{ijk} (EF_{ijk} \times t_{ijk} \times A_{ijk} \times 10^{-6})$$

$$EF_i = EF_c \times SF_w \times SF_p \times SF_o \times SF_{s,r}$$

- EF_{ijk} is the emission factor for rice paddy fields in each category (kg ha^{-1} d^{-1} CH_4).
- T_{ijk} is the number of rice cultivation days in rice paddy fields of each category (day).
- A_{ijk} is the harvest area of rice paddy fields in each category (ha yr^{-1}).
- EF_c is the baseline emission factor (kg ha^{-1} d^{-1} CH_4).
- SF_w is the scale factor used for rice paddy type and water management during the growing season.
- SF_p is the scale factor used for water management before the growing season.
- SF_o is the scale factor used for the application of organic amendments.
- $SF_{s,r}$ is the scale factor used for soil type and rice cultivar.

Our Approach

Table: Regression model test results using 10-Fold CV.

Model	RMSE	MAE	R2
SVR	0.3554±0.0742	0.3305±0.0286	0.6472±0.0334
Decision Tree Regressor (DTR)	0.6694±0.1272	0.5783±0.0600	0.3318±0.0679
Random Forest Regressor (RFR)	0.784±0.1223	0.6299±0.0457	0.2163±0.0338
AdaBoost Regressor	0.1219±0.0114	0.2817±0.0153	0.8769±0.0114
XGBRegressor	0.0477±0.0099	0.1488±0.0124	0.9524±0.0057
MLPRegressor (MLPR)	**0.0091±0.0024**	**0.0739±0.0088**	**0.9905±0.0037**

[b] https://ndma.gov.in/sites/default/files/PDF/FHA/WB_FloodHazardAtlas.pdf

Hyperparameters of MLP regressor model

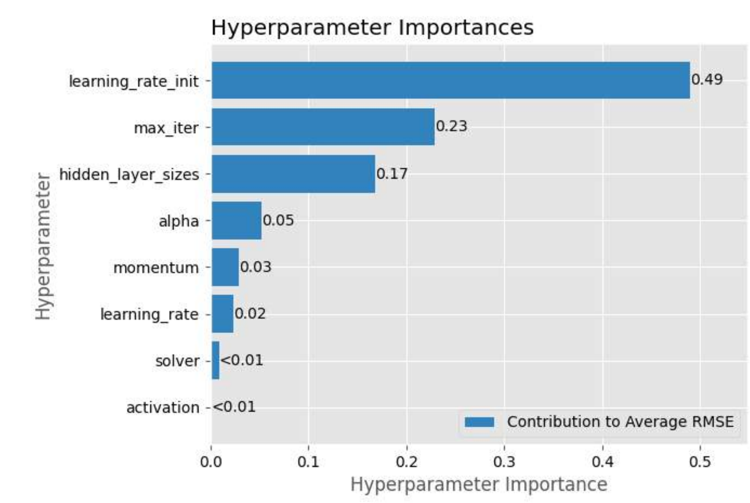

Figure: Importance of Hyperparameters of MLP regressor model

Model	Hyperparameters (HPs)
MLPR	*activation* = 'logistic', *alpha* = 0.000842, *hidden_layer_sizes* = 14, *learning_rate* = 'adaptive', *learning_rate_init* = 0.003694, *max_iter* = 265, *momentum* = 0.071646, *solver* = 'sgd'

Table: MLP regressor model and HPs for regression

SHAP plots of MLP regressor model

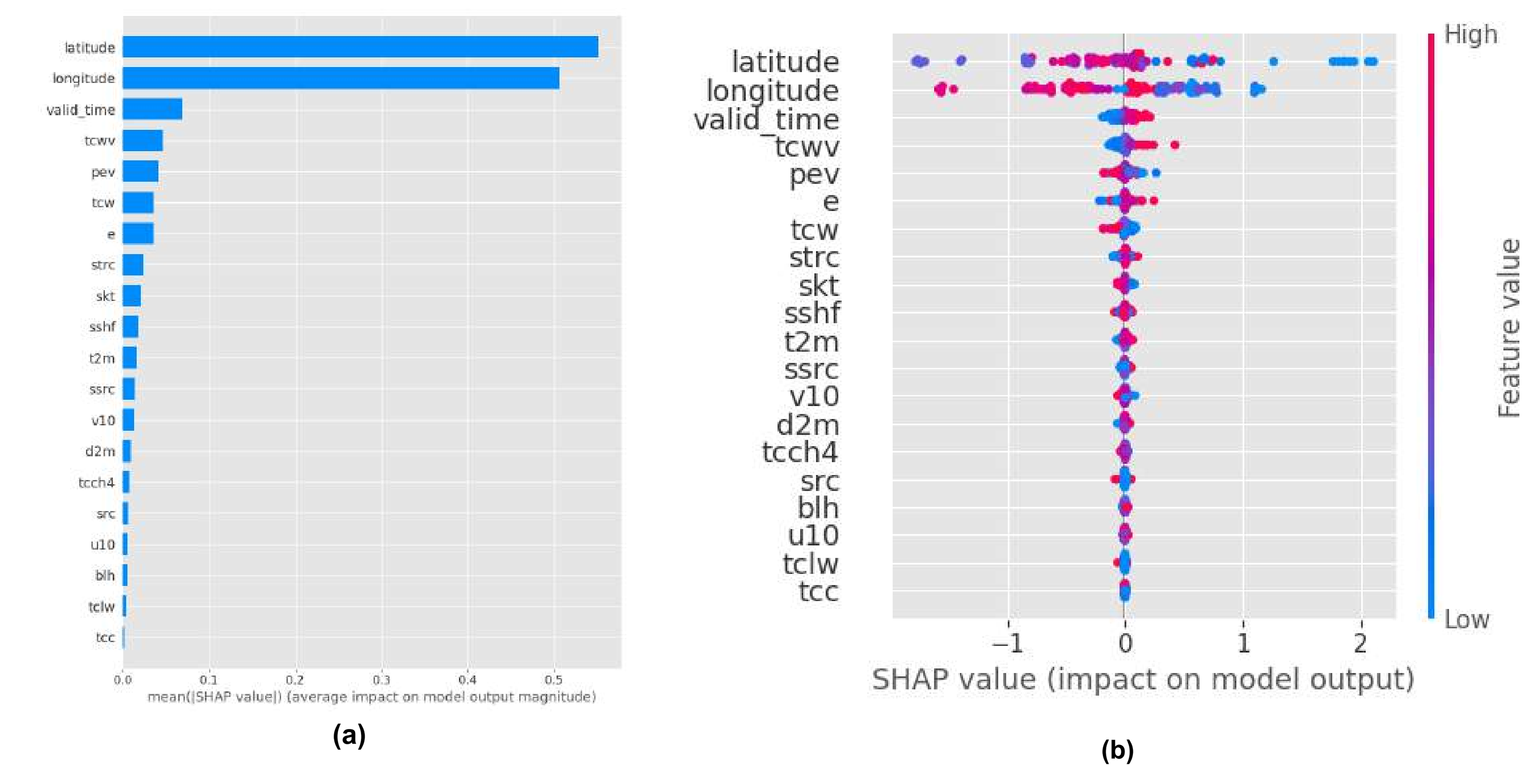

Figure: SHAP plots (a) represents bar plot and (b) represents beeswarm plot

Conclusion

In this work, we propose a framework that combines current research on generating machine learning models for regression tasks to predict methane emissions from rice paddies. Major findings of our work identify which hyperparameters are the most influential in generating ML models and highlight that MLP was the better regressor model across all metrics and RF is the least-performing model for methane prediction (Table 1). Additionally, Shapley values were derived, along with how these values can be used to rank the significance of features on model results.

References

[1] Youngho Kwon et al. "Loss-of-function gs3 allele decreases methane emissions and increases grain yield in rice". In: *Nature Climate Change* 13.12 (2023), pp. 1329–1333.

[2] Swades Pal et al. "Methane emissions only negligibly reduce the ecosystem service value of wetlands and rice paddies in the mature Ganges Delta". In: *Environmental Science and Pollution Research* 29.19 (2022), pp. 27894–27908.