

Article

Why Abundant Biomass Fails to Deliver: Machine Learning Insights into Biogas Production Constraints in Sub-Saharan Africa

Zongrun Song ¹ and Zhiyuan Ma ^{1,2,*}¹ Yantai Institute, China Agricultural University, Yantai 264670, China; songzongrun@cau.edu.cn² National Center for International Research of BioEnergy Science and Technology, China Agricultural University, Beijing 100083, China

* Correspondence: mazy@cau.edu.cn

Abstract

Sub-Saharan Africa is rich in agricultural biomass, yet its biogas utilization is far below its potential. Most earlier studies failed to identify the nonlinear, multi-factor relationships that shape real national biogas yields and fully clarify this imbalance. This study constructs a 2007–2023 panel dataset for ten sub-Saharan African countries, merging agricultural output, socioeconomic, and infrastructure metrics. Gradient Boosting model and SHapley Additive exPlanations (SHAP) analysis are applied for empirical evaluation. SHAP analysis confirms that charcoal consumption yields the largest contribution to biogas production, with a mean absolute SHAP value of 1.018. The correlation between the two variables is negative under the threshold and becomes positive beyond this critical level. Urbanization has an inverted U-shaped correlation with biogas output, and the marginal contributions of predictors vary substantially across sampled countries. Instead, fragile supply chains, rural labor loss, and fierce competition in clean energy markets curb local biogas production. Forecasts show that regional biogas output will continue to fall until 2030. Targeted national policies matching each country's core influencing factors are therefore urgently required.

Keywords: biogas production; machine learning; SHAP; energy poverty

1. Introduction

Energy poverty refers to the lack of adequate and reliable energy services and represents a severe development challenge in the Africa. Millions of people in Africa lack access to modern energy, with sub-Saharan Africa (SSA) being the most severely affected [1]. Energy poverty exacerbates various development difficulties in Africa, while also hindering global sustainable development and increasing climate change risks [2]. The United Nations Sustainable Development Goal (SDG) 7 aims to ensure access to affordable, reliable, sustainable, and modern energy for all [3].

SSA is endowed with abundant energy resources, yet its per capita access to clean energy is the lowest in the world. Currently, over 600 million people in the region lack access to electricity, presenting a stark contrast between resource reserves and actual energy availability [4]. Due to the lack of access to modern energy, rural areas in most sub-Saharan African countries rely heavily on traditional biomass fuels, such as firewood,

Academic Editor: Cinzia Buratti

Received: 11 June 2026

Revised: 15 July 2026

Accepted: 17 July 2026

Published: 18 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

charcoal, and dung cakes. The long-term use of these fuels not only increases the risk of respiratory health problems but also leads to vegetation destruction, soil erosion, and other environmental issues, exacerbating global environmental degradation and climate change [5–7]. With the growth of the population and income, the demand for modern energy in Africa will continue to increase, and energy demand is expected to rise by one-third between 2020 and 2030 [8]. As a clean energy source, biogas can effectively alleviate energy poverty in African countries [9,10]. There are significant differences in the development of the biogas industry across countries worldwide. In Europe and the United States, biogas facilities are large-scale plants. As of 2017, Europe's electricity generation from biogas had reached 64 terawatt-hours, accounting for 70% of the global biogas power generation [11]. In contrast, in Asia and Africa, biogas installations are mostly small-scale household digesters, focusing on meeting household needs such as cooking and lighting [12]. Since 2008, multiple international organizations have supported and funded biogas projects in Africa. Among them, the Africa Biogas Partnership Programme (ABPP) has been implemented in Ethiopia, Kenya, Tanzania, Uganda, and Burkina Faso, resulting in the construction of over 70,000 biogas digesters by 2019 [13]. According to relevant studies, utilizing biomethane generated from African crop residues in combined heat and power (CHP) systems could produce 395 PJ of electricity annually, meeting 16.3% of Africa's electricity demand [14]. If all dairy cattle manure in Africa were processed in biogas digesters and the resulting biogas used for power generation, the electricity generation potential could meet 14% of Africa's total electricity demand, equivalent to 54% of the electricity demand in SSA (excluding South Africa) [15]. Effectively converting biomass resources into energy would actively improve the lives of millions of Africans, thereby helping the continent break free from energy poverty. The suitable tropical climate and agriculture-driven industrial foundation of the SSA region create favorable conditions for the development of biogas technology [16]. In addition to providing energy, biogas also brings a range of benefits, including increasing household income, enhancing environmental performance, reducing labor costs, and improving residents' health [17–19]. Despite the abundant biomass resources, great development potential, and the implementation of several international aid projects in SSA, the overall development level of biogas technology in Africa remains low due to constraints such as economic development levels, technological shortcomings, inadequate policies, and weak infrastructure [20,21].

Regarding the driving factors of biogas development, Kelebe et al. and Berhe et al. approached the issue from the perspective of farmers' adoption behavior and found that household biogas adoption is influenced by multiple factors, including social, demographic, environmental, and access to basic infrastructure. These factors include the age of the household head, household size, education level, number of cattle owned, distance to firewood, and access to electricity, roads, and credit [22,23]. Kamp et al. employed Strategic Niche Management (SNM) and the Multi-Level Perspective (MLP) to analyze the driving factors of biogas development in Ethiopia, such as climate, policies, and resources, as well as bottlenecks, including low purchasing power and technological mismatches [24]. For the analysis and prediction of driving factors in biogas energy, Senol et al. used Artificial Neural Networks (ANNs), autoregressive integrated moving average (ARIMA), and linear regression models to estimate and evaluate the biogas potential from cattle manure in Turkey [25]. Yan et al. calculated the biogas production capacity from agricultural waste in China and explored the correlation between actual rural biogas output and its influencing factors [26]. In contrast, in SSA, relevant research exhibits clear limitations. Most studies focus on micro-level surveys of household adoption behavior or provide qualitative explanations at the macro level. Quantitative cross-country analyses based on national macro-level panel data combined with actual biogas production are relatively rare, and there is insufficient analysis of macro-level driving mechanisms and country-

specific heterogeneity. Furthermore, existing research largely employs traditional linear statistical models for static analysis, without considering the nonlinear interactions among multiple factors. The research content has long remained at the level of estimating theoretical biogas potential and recoverable potential [13–15], failing to integrate realistic conditions such as social, economic, and infrastructural factors to assess and predict achievable biogas potential. Research focusing on sub-Saharan African countries is particularly lacking. Given the reality of abundant biogas resources but low actual utilization rates in this region, existing studies have not been able to provide sufficient explanations through quantitative data analysis. Ihsan Pence et al. and Angelique Mukasine et al. both adopted static cross-validation without temporal testing, SHapley Additive exPlanations (SHAP) decomposition, or medium-to-long term extrapolation [27,28]. Although Zeydin Pala identified the generalization flaws of static validation, it only performed basic feature analysis and failed to construct temporal validation and extrapolation frameworks [29]. Therefore, this study integrates gradient boosting model with SHAP, adopts leave-one-year-out (LOYO) cross-validation, and combines compound annual growth rate (CAGR) and ARIMA dual extrapolation models to establish an integrated analytical framework for biogas driving factors.

This study focuses on sub-Saharan African countries. It employs machine learning models and SHAP to investigate the impacts of multidimensional indicators, including agricultural production, socioeconomic conditions, resources, and infrastructure, on actual biogas production. It aims to reveal the country-specific heterogeneity in the dominant factors of actual biogas production across different sub-Saharan African nations and to forecast biogas production in the near term. The findings will provide quantifiable evidence and scientific support for the differentiated formulation of biogas development strategies in sub-Saharan African countries.

2. Materials and Methods

2.1. Study Area and Data Sources

This study selects ten sub-Saharan African countries: Burkina Faso, Cameroon, Ethiopia, Ghana, Kenya, Nigeria, Rwanda, Senegal, Tanzania, and Uganda. The sample has good regional representativeness and complete biogas data. These agricultural countries have abundant biogas feedstocks, with practice and experience on biogas projects. Covering East, West, and Central Africa with diverse populations and biogas development levels, the sample eliminates regional sampling bias and guarantees reliable results.

This study takes the annual actual biogas production of each country from 2007 to 2023 as the dependent variable, with relevant data sourced from the statistical website of the International Renewable Energy Agency (IRENA). Relatively complete and continuous panel data are publicly available for all selected research countries, which can provide solid data support for empirical analysis. Based on existing research findings and regional development characteristics, an explanatory variable system is constructed comprising 12 indicators selected from three dimensions: agricultural production, socioeconomic conditions, and resources and infrastructure. Biomass is the basic feedstock for biogas, and agricultural outputs reflect the regional raw material supply. Agricultural indicators include crop residues, livestock manure, and total agricultural output. Socioeconomic indicators are urbanization rate, per capita GDP, agricultural labor, and rural population: the first two reflect economic levels and farmers' investment capacity, and the others represent energy demand and labor supply. Resource indicators include rural electricity, clean cooking, charcoal and firewood consumption. Firewood and charcoal use reflect rural reliance on traditional energy, and rural energy facilities directly restrict biogas populariza-

tion. As household income, poverty, and energy infrastructure cannot be directly quantified across countries, we use per capita GDP and urbanization rate to proxy income and poverty, and rural clean cooking and electrification to reflect energy infrastructure.

The original data for all indicators are sourced from publicly available data platforms, such as the Food and Agriculture Organization of the United Nations statistical database (FAOSTAT), the World Bank database, and the United States Department of Agriculture database (USDA). The units and sources of the data used in this study are presented in Table 1.

Table 1. Units and sources of dependent variables and characteristic variables.

Date Type	Unit	Source
Total biogas production	1000 m ³	IRENA [30–37]
Crop residue quantity	kt	FAOSTAT [38]
Livestock manure quantity	kt	FAOSTAT [38]
Total crop output	\$1000, constant 2015 prices	USDA [39]
Total animal output	\$1000, constant 2015 prices	USDA [39]
Urbanization rate	%	World bank [40]
GDP per capita	constant 2015 international \$	World bank [40]
Total agricultural labor	1000 persons	USDA [39]
Rural population	person	World bank [40]
Rural electricity access	% of rural population	World bank [40]
Rural clean cooking access	% of rural population	World bank [40]
Charcoal consumption	TJ	FAOSTAT [38]

Note: Since official data on crop residue and livestock manure volumes are unavailable, this study adopts six major African food and cash crops (maize, millet, rice, sorghum, potato, sweet potato, sugarcane) and six typical livestock species (cattle, sheep, goats, pigs, poultry, horses) for estimation. Calculations of crop residues refer to Table S1, Equations (S1) and (S2). Those for livestock manure are available in Table S2 and Equation (S3). Crop yield and livestock population data are retrieved from FAOSTAT. Crop residue types and residue-to-product ratios are from References [41,42], and daily livestock manure production is from Reference [43].

2.2. Data Preprocessing and Variable Screening

To eliminate the interference of missing original data, dimensional differences, and multicollinearity on model accuracy and stability, this study performed preprocessing steps, including the removal of missing values, numerical feature standardization, and logarithmic transformation of the dependent variable.

To avoid the impact of multicollinearity among independent variables on model stability, the Pearson correlation coefficient method was used for initial feature screening, with $|r| \geq 0.8$ set as the criterion for high collinearity. For groups of variables with high collinearity, the indicator with a stronger correlation with the dependent variable (total biogas production) was retained, whereas redundant variables with weaker correlations were removed.

To further validate the effectiveness of variable selection, the Variance Inflation Factor (VIF) was used for collinearity diagnosis. When $VIF < 10$, there is no significant multicollinearity among the variables, allowing them to be used for model training and analysis.

2.3. Research Framework

2.3.1. Machine Learning Models

To comprehensively compare the biogas prediction performance of various algorithms and enrich the model comparison framework, this study adopted tree-based ensemble machine learning models as the core prediction approaches, with three groups of

benchmark models incorporated to horizontally verify the predictive effectiveness and robustness of the core models. The core tree-based ensemble models consisted of Random Forest, Gradient Boosting Regression (GBR), Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM). Such models exhibit strong capacities for nonlinear fitting, feature interaction capture, and noise resistance, which fit the multidimensional, nonlinear cross-country panel data characterized by intricate variable interactions in this study and ensure favorable prediction accuracy.

The benchmark models were categorized into three groups as follows:

- (1) Classical panel econometric benchmark models, including Pooled Ordinary Least Squares (Pooled OLS) and Fixed Effects (FE) regression;
- (2) Regularized linear models, including Ridge regression and Lasso regression;
- (3) Nonlinear machine learning benchmark models, including Support Vector Regression with Radial Basis Function kernel (SVR, RBF kernel) and K-Nearest Neighbors (KNN), to complete the comparison system.

Since the study sample covers multiple countries and considering the significant inherent differences among countries in terms of resource endowment and economic development foundation, this study employs one-hot encoding to account for country fixed effects. This approach removes the interference of country-specific differences within a unified modeling framework, thereby improving the model accuracy. The model uses five-fold cross-validation for training and evaluation to prevent overfitting and ensure the stability and comparability of results. All the models were implemented in Python 3.9.5 with scikit-learn (v1.6.1), XGBoost (v2.1.4) and LightGBM (v4.6.0).

2.3.2. Model Training and Hyperparameter Tuning

To guarantee the reproducibility of the experiments, the dataset was randomly split into training and test sets at an 8:2 ratio with shuffling. A unified global random seed, SEED = 42, was used in this study. For all machine learning models, a 60-iteration random search algorithm combined with a shuffled 5-fold cross-validation was adopted for automatic hyperparameter optimization. The maximum cross-validation R^2 was used as the optimization objective to screen the optimal hyperparameter combination for each model.

In terms of benchmark econometric models, Pooled OLS and FE regressions contained no tunable hyperparameters and were fitted using ordinary least squares. The regularization coefficient α of the regularized linear models was searched over the log-scale range of 10^{-3} to 10^3 . The hyperparameter search spaces for tree-based ensemble models were set as follows: for Random Forest, $n_estimators$ was selected from {50,100,200}, and max_depth ranged from 3 to 5. Gradient Boosting, XGBoost and LightGBM shared identical basic search ranges: $n_estimators \in \{50,100,200\}$, learning rate between 0.05 and 0.2, and max_depth from 2 to 5. The subsample ratio with candidate values {0.6, 0.8} was only tuned for Gradient Boosting and XGBoost. The additional L1 and L2 regularization terms of XGBoost and LightGBM were tuned within the log interval of 10^{-2} to 10. For SVR, the penalty coefficient C was scanned logarithmically from 10^{-1} to 10^3 , with multiple kernel parameter γ for joint optimization. The number of neighbors for KNN ranged from 3 to 10, and two weighting schemes, uniform and distance, were compared simultaneously.

2.3.3. Model Performance Evaluation

To comprehensively and objectively compare the fitting accuracy and predictive performance of the four machine learning models, several metrics are selected for evaluation: coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and symmetric mean absolute percentage error (SMAPE).

The R^2 value ranges between 0 and 1. The closer the value is to 1, the better the model's fitting effect and the stronger its ability to explain data variability. RMSE and MAE can intuitively reflect the absolute deviation between the model's predicted values and the actual observed values. The lower the indicator values, the higher the numerical fitting accuracy of the model. SMAPE eliminates the influence of unit and magnitude differences in the data, objectively measuring the relative prediction error of the model. The lower the value, the stronger the model's generalization performance. The relevant formulas are shown in Equations (1)–(4). Models with no predictive value and unsatisfactory fitting accuracy are hierarchically eliminated based on the above indicators, while the model with optimal predictive performance is retained for temporal robustness testing.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

$$SMAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2} \times 100\% \quad (4)$$

where y_i and $\hat{y}_i$ are the observed and predicted values of the i -th observation, respectively, $\bar{y}$ is the mean of the observed values, and n is the total number of observations.

2.3.4. Temporal Robustness Test

This study adopted LOYO cross-validation to assess the temporal generalization of the optimal model selected by 5-fold cross-validation. The samples were grouped by observation year: in each iteration, one year's observations served as an independent test set, and the rest for training. Full-year temporal validation yielded averaged fold metrics for evaluating the temporal prediction stability of the model.

Owing to sharp fluctuations in the early data, a segmented analysis was conducted for 2012–2023 (a stable period) to evaluate the model performance under milder data distribution shifts. Significant annual shifts in biogas driving factors increase cross-year extrapolation difficulty, and the 2012–2023 results better reflect the practical long-term predictive capacity of the model in real scenarios.

2.4. Model Interpretability Analysis

Machine learning models commonly suffer from a “black box” problem, where good model fit contrasts with poor interpretability of the specific mechanisms of each variable. Based on the optimal machine learning model, this study introduces the SHAP (v0.49.0) interpretability algorithm. Relying on the Shapley value principle from game theory, it comprehensively explains the non-linear effects, interactions, and country-specific differences in each characteristic variable on biogas production in SSA through global feature importance analysis, partial dependence plot analysis, and cross-country heterogeneity analysis, achieving visual interpretation of the model results.

2.5. Forecast Analysis of Future Production

This study adopts the year-on-year geometric mean growth rate method, namely the CAGR approach, as the primary scheme to extrapolate all explanatory variables. We established an ARIMA extrapolation method for comparison, with the same model order and constraints as the CAGR approach. The two sets of extrapolated variables were fed into the same gradient boosting model for comparison. Equations (S4) and (S5), and the corresponding constraints are detailed in the Supplementary Materials.

Extrapolation rests on steady growth assumptions, yet socioeconomic indicators are disturbed by policies, economic shocks, and technological changes, yielding high predictive uncertainty. Biased socioeconomic extrapolation introduces systematic errors that accumulate annually in long-term biogas projections. We generate 95% confidence intervals via 1000 residual bootstrap resamples to quantify forecast volatility. Full CAGR formulas, ARIMA specifications, and bootstrap procedures are available in Supplementary Material.

3. Results

3.1. Multicollinearity Test and Variable Selection

The Pearson correlation coefficient matrix in Figure 1 shows that some characteristic variables among the 12 initial indicators exhibit high linear correlations. Directly including them in the model would cause multicollinearity issues, compromising model stability. Based on the screening criterion of $|r| \geq 0.8$, for groups of variables with high correlation, redundant variables were removed. Ultimately, six variables were eliminated, and six core explanatory variables were retained: livestock manure quantity, total crop output, urbanization rate, rural clean cooking access rate, rural electricity access rate, and charcoal consumption.

To test the validity of the screened variable set, a VIF test was further conducted. The results are shown in Table 2. The VIF values of each variable range between 4.06 and 8.45, all below the critical value of 10, indicating no significant multicollinearity among the selected core explanatory variables. They are therefore suitable for model training.

Table 2. Variance inflation factor test results of variables

Variable	VIF
Rural electricity access	8.45
Livestock manure quantity	7.59
Urbanization rate	7.11
Charcoal consumption	5.80
Rural clean cooking access	5.72
Total crop output	4.06

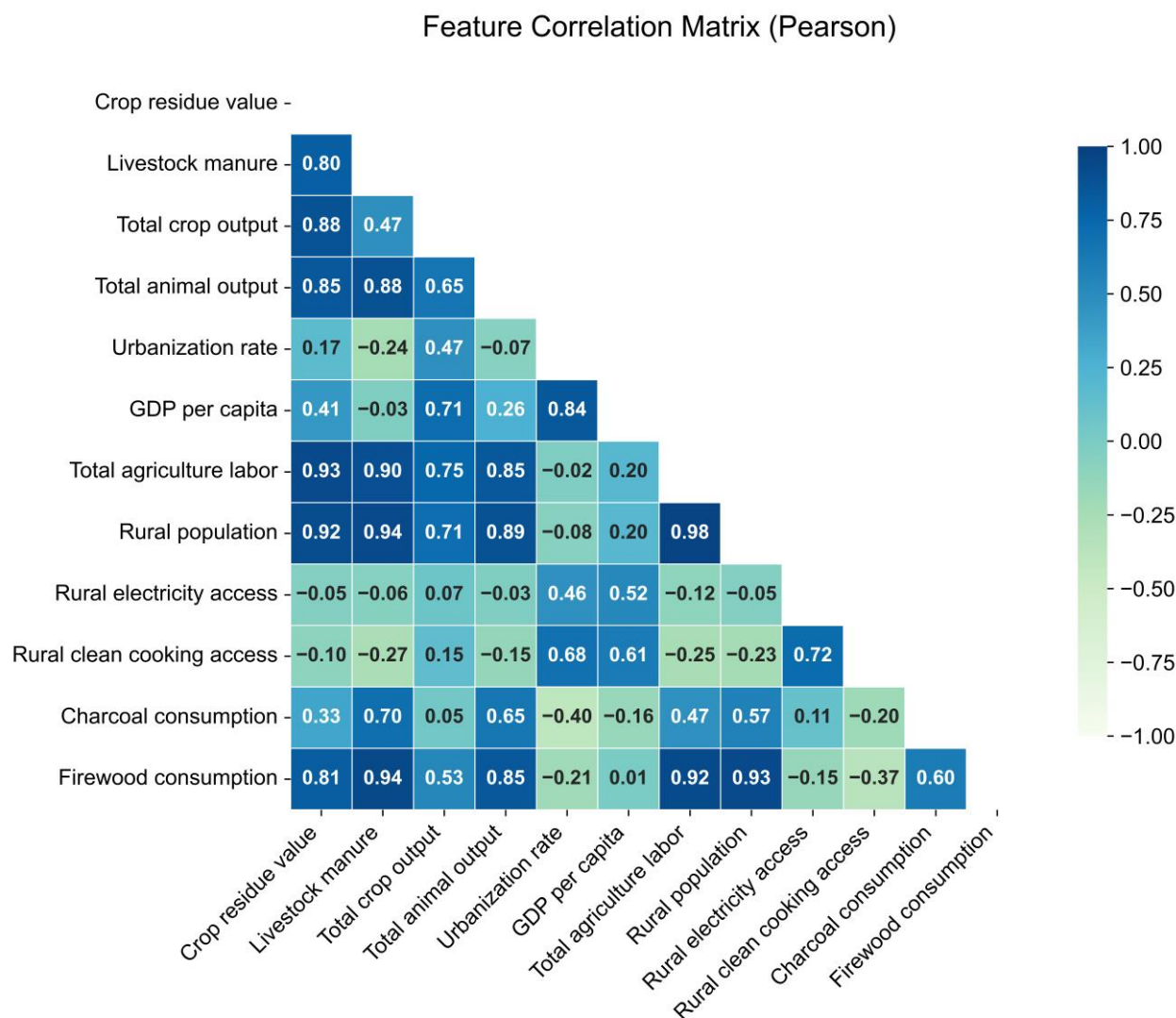


Figure 1. Pearson correlation coefficient matrix of characteristic variables.

3.2. Optimal Model Selection

3.2.1. Multi-Model Performance Comparison

All candidate models in this work were optimized via RandomizedSearchCV with 5-fold cross-validation to eliminate bias from default hyperparameters. The hyperparameter search ranges for each algorithm are listed in Table S3, and the optimal hyperparameter combination of each model was retained to calculate the prediction evaluation indicators.

This study compares the predictive performance of tree-based ensemble models, conventional panel models, regularized linear models, and typical machine learning algorithms (KNN and SVR), with all evaluation metrics summarized in Table 3. Collectively, tree-based ensemble algorithms yield higher prediction accuracy than all alternative modeling strategies, as shown in the density scatter plot in Figure 2.

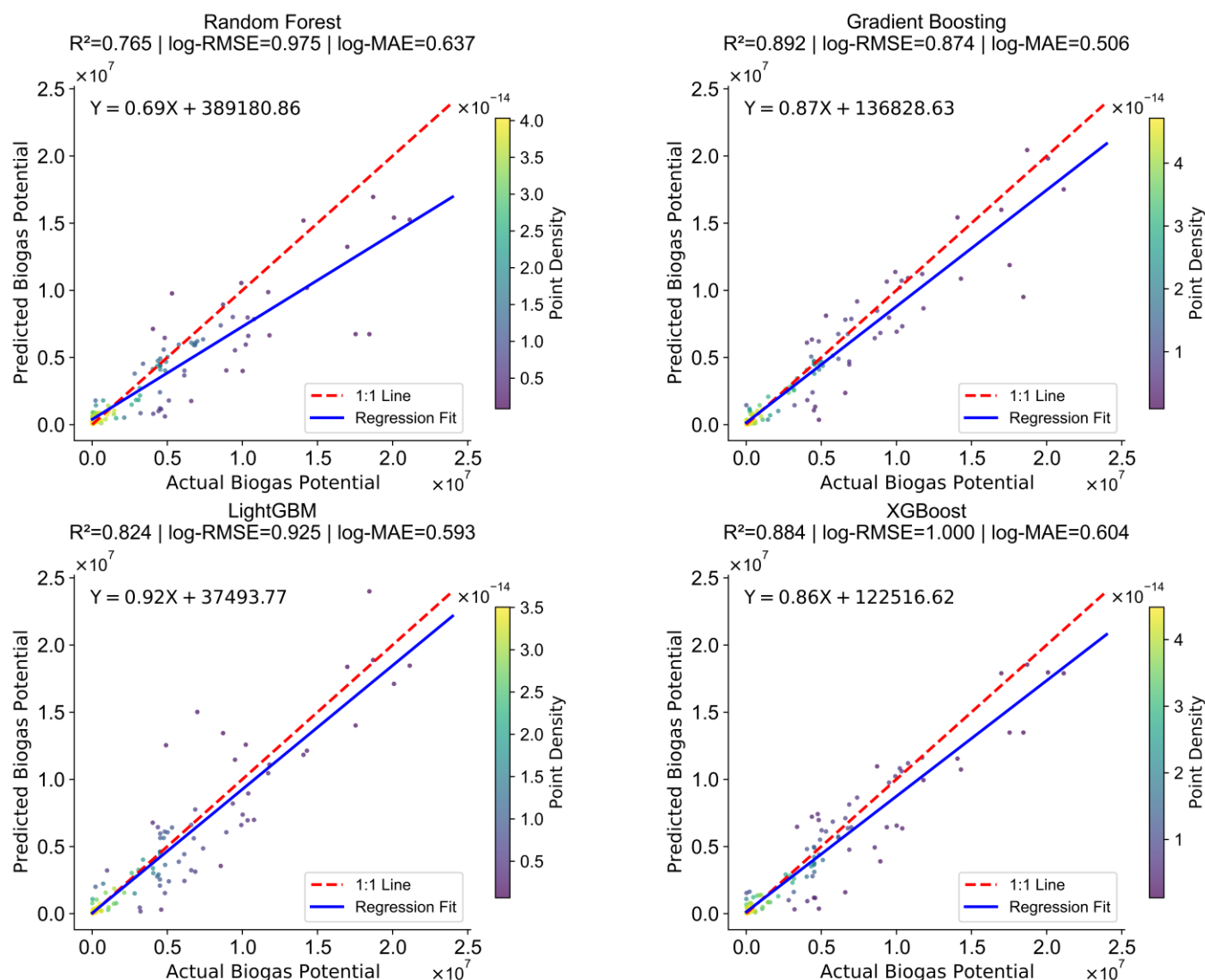


Figure 2. Scatter density plot of biogas fitting for four tree-based machine learning models.

Gradient Boosting model presents the optimal performance, with an average R^2 of 0.892 and the minimum $\log\text{-RMSE}$ and $\log\text{-MAE}$ among all candidate models. XGBoost and LightGBM show the second and third highest R^2 , respectively, whereas Random Forest performs the worst within the tree ensemble group. Meanwhile, the Gradient Boosting model achieves an SMAPE of 1.21%, indicating an extremely low overall prediction error. Relying on layer-wise iterative fitting strategies, Gradient Boosting can identify complex nonlinear variable interactions, yielding superior fitting accuracy and generalization stability.

In contrast, other model categories have conspicuous defects. Pooled OLS yields an R^2 of 0.096, while the fixed-effects model produces a negative R^2 . Additionally, both models feature extremely large standard deviations, which demonstrate that linear panel models fail to capture the nonlinear characteristics of the dataset. Constrained by linear assumptions, Ridge and Lasso regressions only achieve an average R^2 of 0.6, with limited explanatory power for the dependent variable. The R^2 values of KNN and SVR are both below 0.8, accompanied by higher $\log\text{-RMSE}$ and poor generalization stability, indicating inferior overall predictive performance compared with Gradient Boosting model. Specifically, KNN relies on local samples for prediction and cannot effectively identify global

data patterns, whereas SVR is highly sensitive to hyperparameter configurations and outliers, resulting in a much lower fitting ceiling. After a comprehensive comparison, Gradient Boosting model is selected as the optimal prediction model in this study.

Table 3. Comparison of predictive performance for different models.

Model Category	Model	R ²	RMSE (log)	MAE (log)
Tree-based Ensemble Models	Gradient Boosting	0.892 ± 0.072	0.874	0.506
	LightGBM	0.824 ± 0.087	0.925	0.593
	XGBoost	0.884 ± 0.016	1.000	0.604
	Random Forest	0.765 ± 0.161	0.975	0.637
Traditional Panel Models	Pooled OLS	0.096 ± 1.054	1.390	0.984
	Fixed Effects Model	−3.130 ± 5.087	1.163	0.901
Regularized Linear Models	Ridge	0.664 ± 0.071	1.266	0.911
	Lasso	0.654 ± 0.101	1.318	0.949
Other Machine Learning Models	SVR (RBF kernel)	0.765 ± 0.164	1.152	0.751
	KNN	0.790 ± 0.250	0.923	0.562

3.2.2. Model Robustness Analysis

The LOYO cross-validation method was adopted to test the robustness and stability of the Gradient Boosting time-series model. The validation results indicate that the global average LOYO R² from 2007 to 2023 was −13.391, revealing poor overall robustness and generalization ability of the Gradient Boosting model. The negative R² value was primarily caused by time-series noise, abnormal driving factors, and inter-annual sample heterogeneity in the early stage rather than model fitting failure. After eliminating disturbed data before the turning point, the segmented average LOYO R² reached 0.738 during 2012–2023, with a standard deviation of merely 0.324. This demonstrates that the Gradient Boosting model achieves favorable fitting accuracy with slight annual metric fluctuations and strong anti-interference performance in this period. In general, the Gradient Boosting model presents obvious temporal heterogeneity in robustness. Although it performs unstably in full-time-series modeling, it possesses satisfactory stability and reliability after 2012, which can meet the requirements of subsequent time-series simulation and predictive factor analysis.

3.3. Analysis of Influencing Mechanisms Based on SHAP

3.3.1. Global Feature Importance Analysis

Interpretability analysis of the optimal Gradient Boosting model using SHAP was conducted to explore the influence intensity of each feature value on actual biogas production. The results are shown in Figure 3. There were significant stepwise differences in the influence intensity of the six core variables on biogas production. Charcoal consumption had the most significant impact on actual biogas production, with a mean absolute SHAP value of 1.018. Urbanization rate, total crop output, and rural electricity access exerted the second-largest predictive impacts, with mean absolute SHAP values of 0.458, 0.361, and 0.337, respectively. In contrast, livestock manure quantity and rural clean cooking access had relatively weaker effects on actual biogas production, with limited contributions. As visualized in the global SHAP dependence plots of Figure 4, charcoal consumption, rural electricity access, livestock manure quantity, and rural clean cooking access displayed identical low-range response patterns. Their SHAP values gradually increased from negative to positive as feature values grew, implying a shift in marginal predictive correlation with biogas production. By comparison, urbanization rate and total crop output showed distinct threshold behavior: after reaching a certain critical level,

higher values corresponded to markedly negative SHAP values, indicating a negative marginal linkage with biogas output under high feature magnitudes.

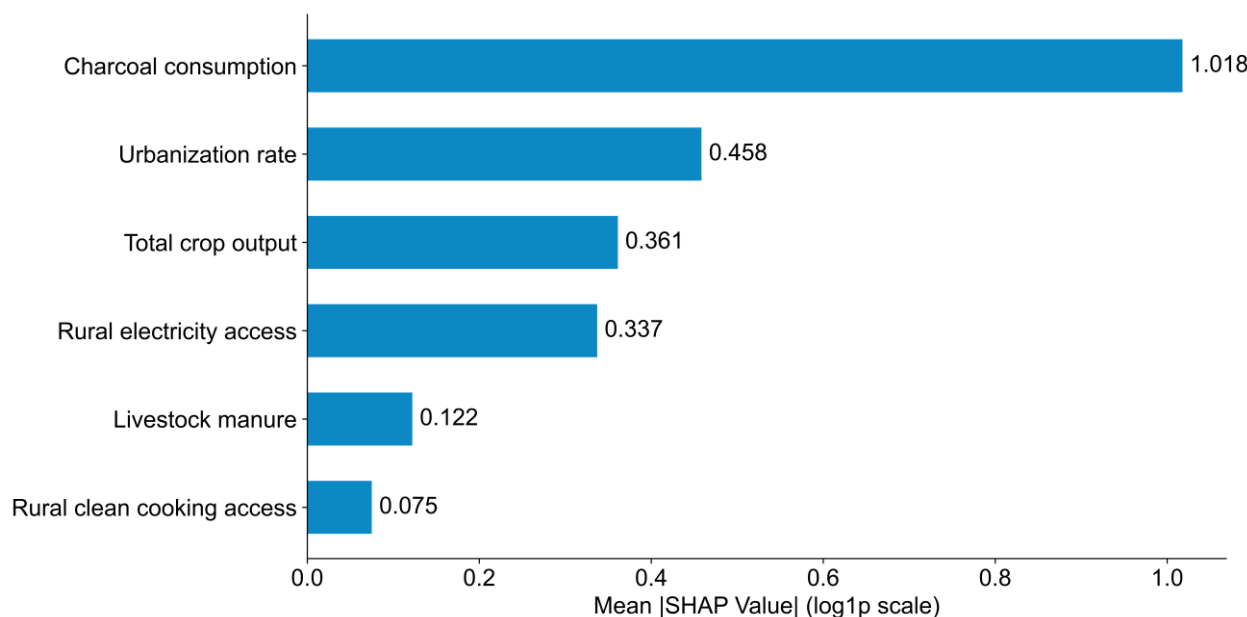


Figure 3. Global feature importance ranking for biogas production.

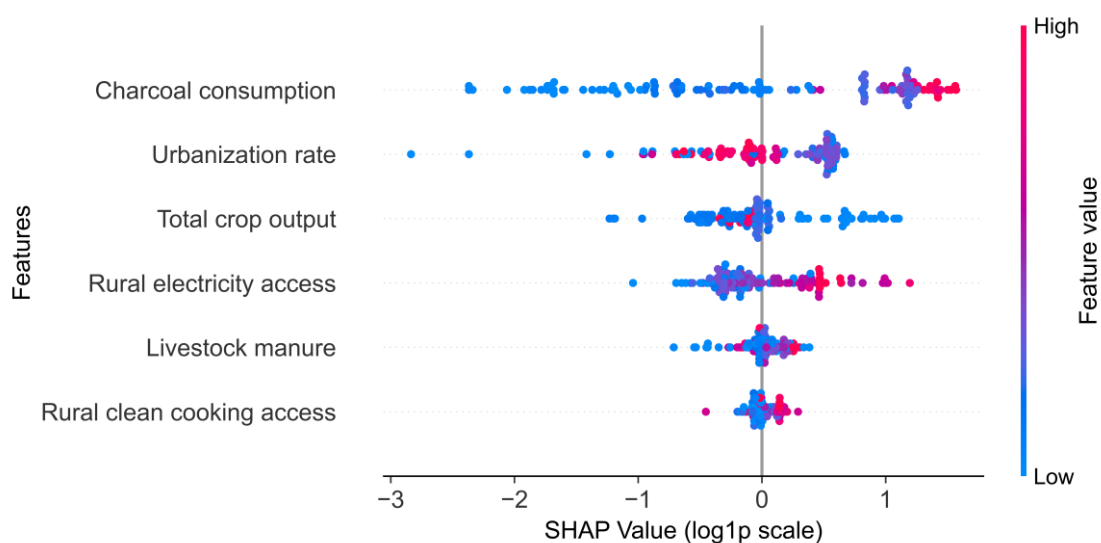


Figure 4. SHAP beeswarm plot for global features of biogas production.

3.3.2. Marginal Effect of Single Variable

The SHAP dependence plots in Figure 5 explicitly characterize the marginal effect dynamics and critical threshold features of each explanatory core variable.

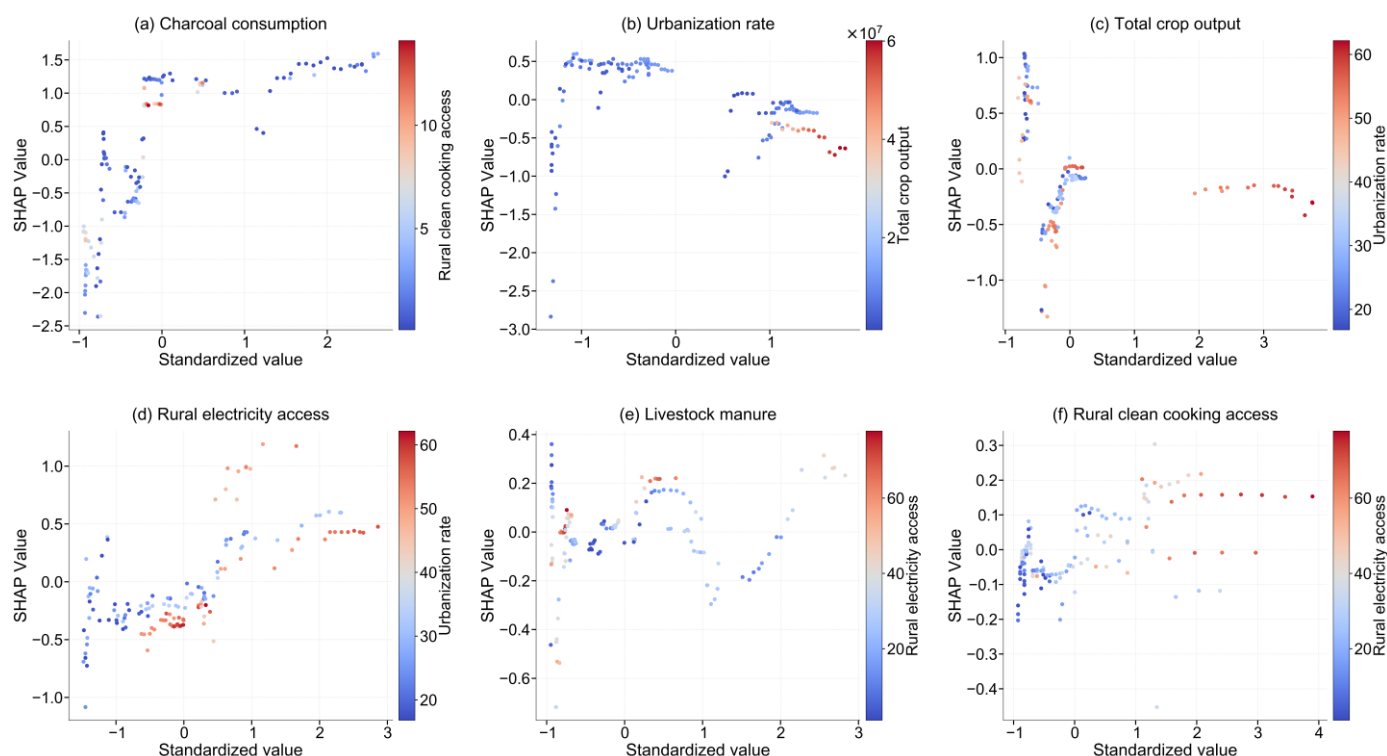


Figure 5. SHAP dependence plots for core influencing factors.

Charcoal consumption exhibits obvious threshold segmentation characteristics. When its standardized value is less than zero, this factor persistently yields negative marginal predictive associations and corresponds to lower predicted biogas output. After exceeding the critical threshold, charcoal consumption maintains positive SHAP contributions and greatly increases the predicted biogas yield.

The marginal predictive contribution of the urbanization rate shows a typical inverted U-shaped trend. At the preliminary urbanization stage, the rising proportion of the urban population continuously boosts positive SHAP values; after crossing the inflection point, its predictive contribution turns negative, and the magnitude of this negative marginal association becomes increasingly prominent.

Rural electricity access rate presents a monotonically rising marginal effect. Low electricity coverage has obvious negative marginal predictive associations, whereas high coverage yields persistent and stable positive predictive contributions.

Standardized livestock manure exhibits mostly negative SHAP values below the zero threshold. After exceeding this threshold, its predictive contribution temporarily rises to positive, then declines back to negative as standardized values increase moderately, and rebounds positively again at high levels.

Access to rural clean cooking exhibits negative marginal predictive associations at low levels and positive predictive contributions at high levels. Nevertheless, the overall intensity of the influence of this variable, whether negative or positive, remains relatively weak.

3.3.3. Interactive Effects Between Variables

The color gradients in the SHAP dependence plots (Figure 5) visually demonstrate the interactive relationships across different feature pairs.

A positive buffering interaction exists between charcoal consumption and rural clean cooking access. Low charcoal consumption corresponds to negative SHAP values and lower biogas predictions. However, the popularization of clean cooking facilities can

weaken such negative marginal correlations. Areas with high charcoal consumption generally lack sufficient clean cooking infrastructure, which matches the intervals of maximum predicted biogas potential.

For observations with high urbanization levels, higher total crop output corresponds to lower SHAP values. This finding reveals that high crop output coupled with excessive urbanization jointly generates combined downward marginal contributions to predicted biogas yields. The two variables exhibit a synergistic negative interaction within high-value intervals.

At low rural electricity access levels, samples with higher urbanization show stronger negative SHAP values, whereas highly urbanized samples produce substantial positive SHAP contributions when rural power coverage is elevated, indicating that complete rural electricity infrastructure serves as the key to counteracting the adverse impacts of excessive urbanization on biogas yield.

Rural electricity access rate has a bidirectional positive moderating interaction with livestock manure. Improved rural electricity coverage can not only alleviate the negative effects caused by insufficient livestock manure supply, but also strengthen the positive predictive contributions from abundant livestock waste resources.

A weak positive interaction was detected between rural clean cooking access and rural electricity access rate.

3.3.4. Cross-Country Heterogeneity Analysis

A cross-country feature importance analysis was conducted based on the prediction model constructed using actual biogas production data from ten sub-Saharan African countries. The results indicate significant cross-country heterogeneity in biogas production patterns among the ten countries (Figure 6).

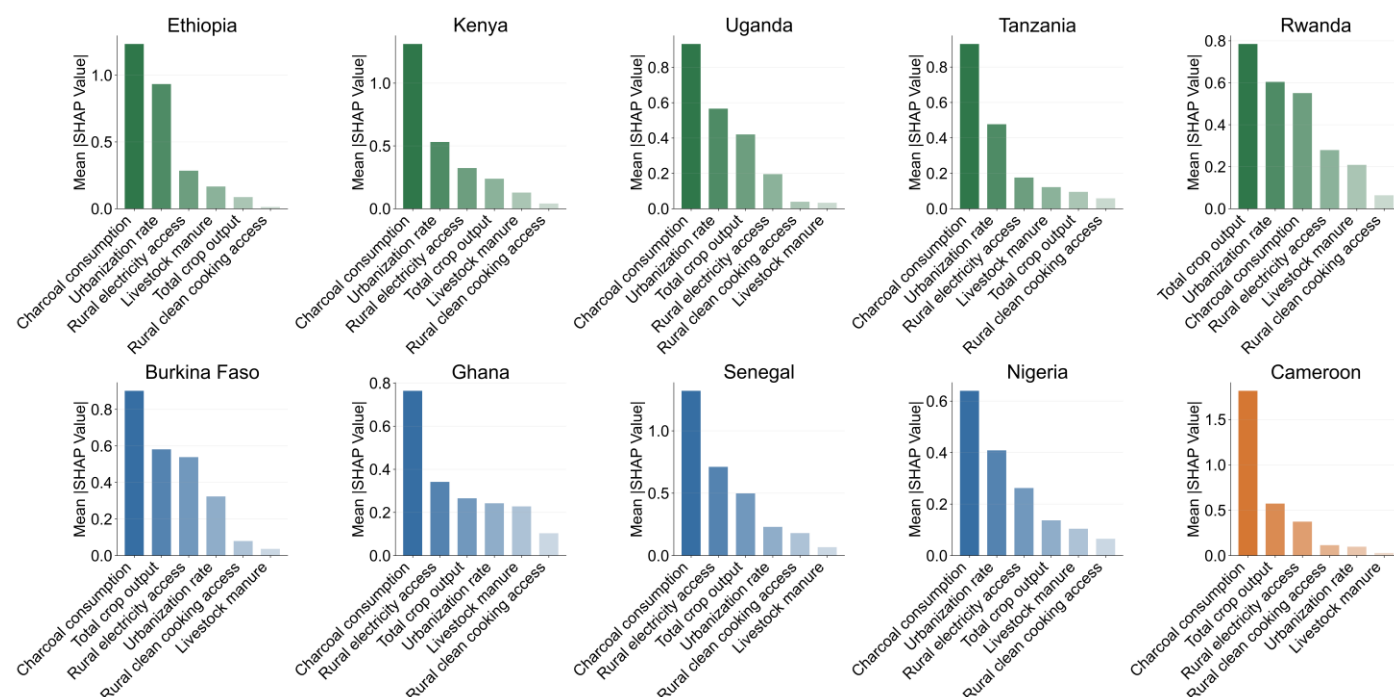


Figure 6. Feature importance for biogas production in ten sub-Saharan African countries.

Charcoal consumption was the dominant predictive determinant of actual biogas production in nine countries, except Rwanda, with mean absolute SHAP values ranging from 0.5 to 1.4 across these nations. It serves as the core predictor of biogas development

in most countries in this region, while the total crop output was the dominant factor shaping biogas production in Rwanda.

The influence intensity of each secondary factor presents prominent country-specific disparities. Urbanization rate functions as a secondary influencing factor in six countries: Ethiopia, Kenya, Uganda, Tanzania, Rwanda, and Nigeria. Total crop output constitutes a secondary predictive factor for biogas production in Burkina Faso and Cameroon, while rural electricity access exhibits comparable secondary predictive importance in Ghana and Senegal. Overall, rural clean cooking access rate and livestock manure quantity are low-to-moderate contributing variables in all surveyed countries, whose mean absolute SHAP values are generally less than 0.3, indicating their limited overall predictive power. Among all samples, livestock manure quantity ranks the lowest in feature importance for Uganda, Burkina Faso, Senegal, and Cameroon. For the remaining countries, rural clean cooking access rate exhibits the minimum feature importance, making it the weakest predictor to interpret variations in biogas output.

3.4. Temporal Evolution and Forecast Analysis of Biogas Production

Figure 7 shows the temporal evolution and annual growth rate changes in total biogas production in the SSA region. From 2007 to 2010, the African biogas sector was in its initial development stage, with a low base but significant growth rate and rapid overall expansion. From 2011 to 2015, the growth rate of biogas production slowed down, but total production remained on a high-growth trajectory. After 2015, the growth rate declined substantially and stabilized between 6% and 11%, indicating weakened growth momentum, with the first negative growth observed in 2023. The forecast results of the CAGR method indicate that after 2024, the total regional biogas production will gradually decrease, with the annual growth rate also turning from positive to negative, reaching $5.17 \times 10^7 \text{ m}^3$ by 2030. According to IRENA statistics, the biogas production of the ten sub-Saharan African countries in 2024 was $6.60 \times 10^7 \text{ m}^3$, showing a decline compared to 2023. The relative error between the model's predicted biogas production for 2024 and the actual data was 0.28%, indicating the high reliability of the model's predictions. In contrast, the forecasted biogas production for the future based on the ARIMA method was consistently lower than that of the CAGR method (Figure S1). Specifically, in 2024, the biogas production predicted by the ARIMA method was about $6.42 \times 10^7 \text{ m}^3$, which is 3.01% lower than the CAGR prediction result and 2.74% lower than the actual data released by IRENA. By 2030, the ARIMA method predicts the total regional biogas production to drop to $4.58 \times 10^7 \text{ m}^3$, which is 11.49% lower than the CAGR prediction result ($5.17 \times 10^7 \text{ m}^3$), and the decline rate of the ARIMA prediction is significantly higher than that of the CAGR method. The 95% confidence interval width for the predictions was generally within a reasonable range, approximately 10% of the predicted value, suggesting controllable uncertainty and good reliability of the results.

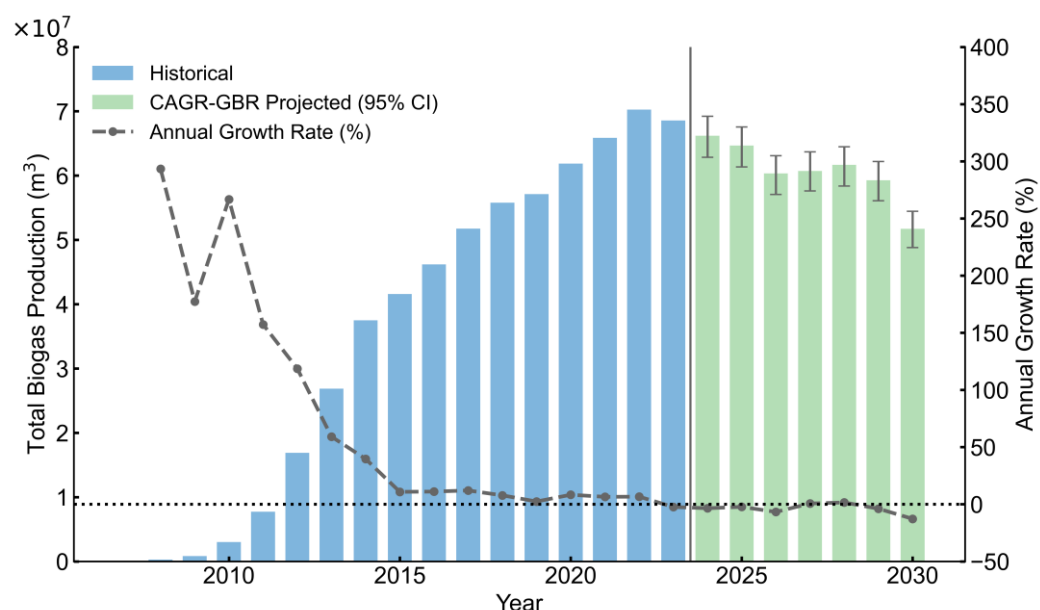


Figure 7. Historical total biogas production in SSA and CAGR-based GBR projections, 2024–2030.

4. Discussion

Global SHAP interpretability analysis showed that charcoal consumption was the primary predictive determinant of variations in biogas yield, with a mean absolute SHAP value of 1.018. The predictive attribution magnitude of the model's predictions of biogas production was far greater than that of the other variables. SHAP dependence analysis indicated that low charcoal consumption was associated with lower predicted biogas production, whereas high charcoal consumption was associated with higher predicted values. There was a significant and strong interaction effect between charcoal consumption and rural clean cooking access. Among regions with limited clean cooking coverage, low charcoal consumption is associated with smaller predicted biogas values. Most samples with medium and high charcoal consumption came from areas with low rural access to clean cooking, where charcoal consumption delivered its maximum positive marginal attribution to biogas potential. Few observations simultaneously show high charcoal consumption and widespread clean cooking access, which indicates two mutually exclusive pathways for rural energy development. Regions with low charcoal consumption and inadequate clean cooking facilities feature weak household energy demand and biogas adoption. In contrast, areas with heavy charcoal reliance and insufficient clean energy infrastructure form a strong demand for energy substitution, corresponding to a higher predicted biogas production potential.

In addition to the core explanatory predictor, other variables also make heterogeneous marginal contributions to the model's predictions of biogas production. Urbanization rate exhibited an inverted U-shaped functional association with predicted biogas production. In the early stages of urban development, the predicted biogas production remains low. In Africa, residents in rural areas normally cannot afford a standard household biogas system, with an upfront cost of \$500–800, which is far higher than that of other clean cooking devices. Although household biogas is efficient and pays off within months, the full upfront cost of a biogas stove and digester totals six times the local households' monthly income, compared with merely half a month's income for a Liquefied Petroleum Gas (LPG) stove. This heavy upfront expense discourages households from adopting biogas. Moderate urbanization corresponded to higher predictions, whereas excessively high urbanization corresponded to lower predictions. In moderate urbanization stages, this pattern may be accompanied by capital inflow, technology spillover, and upgrading of

energy services, which may favor biogas development, whereas excessive urbanization may be accompanied by labor shortages, limiting the routine maintenance of household biogas digesters [44]. SHAP analysis revealed an interaction effect between crop output and urbanization rate: higher crop output was consistently associated with lower predicted biogas production, possibly because, under urbanization pressure, more crop residues are diverted to livestock feeding and commercial processing, reducing the feedstock available for biogas fermentation. The SHAP results for rural electricity access show a nonlinear pattern: at the initial stage of electrification, the predicted biogas production is low, but after the access rate surpasses a certain activation threshold, the predicted production rises with increasing access, suggesting a possible complementary predictive relationship between improved electricity infrastructure and biogas development. Livestock manure showed relatively low SHAP predictive importance for biogas production, consistent with the documented inefficiencies in manure collection, storage, and transportation in the SSA region [45]. Rural access to clean cooking energy had the lowest SHAP importance across all explanatory variables. Notably, SHAP interaction effects revealed that the predictive contribution of livestock manure to biogas production increased alongside higher rural electricity access, implying that improved power infrastructure can strengthen the predictive relevance of manure within the model. Meanwhile, the interactive patterns of rural electricity access and urbanization rate illustrate that insufficient power supply amplifies the negative statistical correlation linking high urbanization to biogas output. For the interaction between rural clean cooking access and rural electricity access, most information embedded in the clean cooking variable overlaps with that conveyed by the electricity coverage indicator.

In summary, the SHAP analysis indicated that crop yields and livestock manure had low predictive importance in the model, suggesting that, based on the statistical patterns of this model alone, feedstock availability was not the core factor constraining biogas output. The model-derived associations indicate two possible restrictive mechanisms. First, inadequate supporting infrastructure for biomass utilization may prevent feedstocks from being fully converted into practical biogas production capacity. Second, the association patterns of high urbanization with lower predicted biogas production could be explained by labor shortages, insufficient digester maintenance, and inter-fuel competition. Combined with the declining predictive contribution of charcoal consumption, these associations imply that weakening energy substitution demand and intensifying clean energy competition may impede the large-scale expansion of the biogas industry.

Affected by the above constraints, the overall development quality of the biogas industry in sub-Saharan Africa is unsatisfactory, and the industry exhibits prominent instability. From 2007 to 2023, driven by local biogas projects and international aid, biogas output in sub-Saharan Africa achieved rapid growth in the initial stage, followed by a slowdown. Although the total installed capacity of biogas digesters has continued to rise, many early built facilities have been abandoned, resulting in poor long-term operational stability of the industry. The abandonment rate of biogas digesters in Kenya, Tanzania, and Uganda has reached 27% [46]. All explanatory variables in this model were forecasted via historical time series extrapolation, resting on the strict assumption of stable long-term trends. Biogas industry development is vulnerable to sudden energy policy adjustments, regional economic slumps, and affordable clean energy innovations, any of which may alter historical growth patterns. Model predictions based on all predictive drivers indicate that the growth rate of regional biogas output will approach zero, with the total output on a declining trend and rising risks of industry shrinkage. It is estimated that biogas output in 2030 will decrease by approximately 16.8 million cubic meters compared with 2022, a volume sufficient to meet the daily cooking gas demand of 30,000 five-person households [47]. Extrapolation brings notable uncertainty to predictions: future biogas subsidies

or economic downturns cutting household investment in digesters could create large gaps between the real 2030 biogas output and our baseline forecast. Random shifts in policies and macroeconomics are the main causes of prediction errors, which may reduce or worsen the projected industrial decline risks. Constrained by overlapping adverse factors, the regional biogas industry suffers from insufficient growth momentum and continuous scale contraction, calling for targeted policy guidance and support.

Given the differences in the leading predictors of biogas production across countries, each nation should accurately identify its own development pain points and formulate targeted development strategies. For the nine countries dominated by charcoal consumption, including Ethiopia, Kenya, and Tanzania, governments should scale up biogas facilities to meet energy substitution demands and advance a steady energy transition to offset the adverse impacts of shrinking charcoal consumption on the biogas sector. Total crop output was the leading predictor in Rwanda. The country should optimize the allocation of crop straw resources, mitigate conflicts over competing biomass resources, and increase the conversion efficiency of crop straw into biogas. Urbanization served as a secondary driving factor in most countries studied. The biogas industry can take advantage of the capital and technological dividends brought about by urbanization to advance industrial standardization and large-scale deployment. Meanwhile, policymakers need to tackle challenges such as labor outflow and inadequate operation and maintenance of biogas facilities, stemming from excessive urban sprawl. For countries where rural electricity access is a secondary predictor, improving the supporting power infrastructure in rural areas can effectively expand biogas production capacity. Furthermore, empirical evidence reveals that the overall clean energy market landscape across the region hinders biogas industry upgrading. Rural clean cooking penetration ranks the lowest in predictive importance for biogas output in the six sampled countries (Ethiopia, Kenya, Tanzania, Rwanda, Ghana, and Nigeria), indicating that biogas faces competitive disadvantages within the local clean energy mix. These countries must strengthen the competitiveness of their biogas industries to prevent further erosion of market share by alternative energy sources, such as LPG. In the remaining countries, livestock manure had the weakest predictive importance; authorities should enhance manure collection systems to facilitate its conversion into biogas energy. However, the above policy implications rely on ten regional countries and model assumptions, with limited generalizability. Moving forward, each country can adopt differentiated policy portfolios tailored to its dominant predictors to break through industrial constraints, fully unlock local biomass resource potential, alleviate rural energy poverty, and advance the achievement of the United Nations Sustainable Development Goal 7. This study constructs a differentiated framework for biogas development. That is consistent with recent African energy research. Existing studies oppose uniform, one-size-fits-all policies. Instead, they recommend targeted strategies tailored to countries that depend on charcoal, crop waste, or limited livestock resources. Previous studies have overly relied on short-term donor aid and subsidies. Mutahi and Roopnarain et al. suggested boosting biogas dissemination via technical training and outreach [15,48], while Getaneh et al. argued that raising household income can also facilitate biogas adoption [49]. We extend targeted policy recommendations for country groups with distinct resource and energy usage characteristics, covering charcoal-dependent economies, crop-dominant countries, and regions with low penetration of clean rural cooking fuels.

This study had several limitations. First, the research only covers 10 countries in sub-Saharan Africa, and missing statistical data further reduce the generalizability of the conclusions. Second, due to the difficulty of quantifying core institutional and economic factors such as subsidies, financing access, institutional quality, and energy prices, the variable system used in this study is somewhat limited. In addition, unpredictable policy

changes across countries may cause deviations in biogas output forecasts. Third, the SHAP analysis can only identify correlations between variables, rather than rigorous causal relationships. Fourth, the Gradient Boosting model exhibits prominent temporal heterogeneity. Temporal robustness tests reveal weakened predictive stability across 2007–2023; reliable forecasts are confined to the steady 2012–2023 subperiod. The model generates credible temporal predictions only for later steady periods and fails to maintain consistent performance over the full temporal span.

Future research can be expanded in four aspects. First, we should expand the sample of cross-country studies and conduct an in-depth analysis of regional heterogeneity. Second, quantitative subsidies, financing, and institutional and infrastructure indicators should be introduced to distinguish the differential impacts of resource endowments and multi-dimensional policy interventions on the biogas industry and improve the rationality of variable selection. Third, micro-survey data should be combined to conduct case studies on typical countries to deepen the understanding of the driving mechanisms of the biogas industry. Fourth, time-series modeling frameworks can be optimized to mitigate performance fluctuations caused by annual data distribution shifts and expand the time range for stable predictions.

5. Conclusions

This study applied Gradient Boosting and SHAP analysis to a 2007–2023 panel dataset of ten sub-Saharan African countries to quantitatively examine the predictive determinants, cross-country heterogeneity, and near-term trends of actual biogas production. We demonstrated that charcoal consumption, rather than biomass resource endowments, is the dominant predictor of biogas production, with its association following a threshold-dependent functional pattern that coincides with the dynamics of rural energy substitution. We further showed that the constraints on biogas development in the region arise primarily from inadequate biomass supply chains, weak infrastructure, rural labor outflow, and clean energy competition, rather than from raw material scarcity.

The empirical findings have direct implications for energy policy in sub-Saharan Africa. The pronounced cross-country heterogeneity in predictive association patterns points to the inadequacy of uniform regional strategies and calls for country-specific interventions aligned with each nation's top predictive variables. Countries where charcoal consumption is the primary predictive determinant should link biogas expansion to energy substitution demand. Those featuring crop output as the primary predictor should prioritize residue allocation efficiency. Four directions merit further investigation: optimizing the time-series framework to broaden its reliable temporal scope, expanding the country sample to improve external validity, incorporating quantitative policy indicators to disentangle the effects of resource endowments and institutional interventions, and integrating micro-level survey data to advance mechanistic analysis of household-level dynamics governing biogas adoption and abandonment.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/su18147365/s1>, Table S1. Key parameters of crop residues; Table S2. Key parameters of livestock manure; Figure S1. Comparison of Future Biogas Predictions between CAGR and ARIMA Methods; Table S3. Optimal hyperparameter combinations for all models.

Author Contributions: Z.S., conceptualization, methodology, software and writing—original draft preparation; Z.M., writing—review and editing, visualization, supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Yantai Integrated University-City Development Project (2023XDRHXMPT12) and (2024XDRHXMPT13).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Mukhtar, M.; Adun, H.; Cai, D.; Obiora, S.; Taiwo, M.; Ni, T.; Ozsahin, D.U.; Bamisile, O. Juxtaposing Sub-Sahara Africa's Energy Poverty and Renewable Energy Potential. *Sci. Rep.* **2023**, *13*, 11643. <https://doi.org/10.1038/s41598-023-38642-4>.
2. Filho, W.L.; Gatto, A.; Sharifi, A.; Salvia, A.L.; Guevara, Z.; Awoniyi, S.; Mang-Benza, C.; Nwedu, C.N.; Surroop, D.; Teddy, K.O.; et al. Energy Poverty in African Countries: An Assessment of Trends and Policies. *Energy Res. Soc. Sci.* **2024**, *117*, 103664. <https://doi.org/10.1016/j.erss.2024.103664>.
3. Goal 7 | Department of Economic and Social Affairs. Available online: <https://sdgs.un.org/goals/goal7> (accessed on 22 May 2026).
4. Tagne, R.F.T.; Dong, X.; Anagho, S.G.; Kaiser, S.; Ulgiati, S. Technologies, Challenges and Perspectives of Biogas Production within an Agricultural Context. The Case of China and Africa. *Environ. Dev. Sustain.* **2021**, *23*, 14799–14826. <https://doi.org/10.1007/s10668-021-01272-9>.
5. Derebe, B.; Alemu, A.; Asfaw, Z. Fuelwood Dependence and Alternative Energy Sources in Ethiopia: A Systematic Review. *Discov. Sustain.* **2025**, *6*, 99. <https://doi.org/10.1007/s43621-024-00721-5>.
6. Sibanda, T.; Uzabakirho, J.D. Animal Manure as an Alternative Bioenergy Resource in Rural Sub-Saharan Africa: Present Insights, Challenges, and Prospects for Future Advancements. *Energies* **2024**, *17*, 1839. <https://doi.org/10.3390/en17081839>.
7. Wassie, Y.T.; Adaramola, M.S. Analysing Household Biogas Utilization and Impact in Rural Ethiopia: Lessons and Policy Implications for Sub-Saharan Africa. *Sci. Afr.* **2020**, *9*, e00474. <https://doi.org/10.1016/j.sciaf.2020.e00474>.
8. IEA Africa Energy Outlook 2022. Available online: <https://www.iea.org/reports/africa-energy-outlook-2022> (accessed on 23 May 2026).
9. Mmusi, K.; Mudiwa, J.; Rakgati, E.; Vishwanathan, V. Biogas a Sustainable Source of Clean Energy in Sub Saharan Africa: Challenges and Opportunities. *J. App. Mat. Sci. Eng. Res.* **2021**, *5*, 7–12.
10. Surendra, K.C.; Takara, D.; Hashimoto, A.G.; Khanal, S.K. Biogas as a Sustainable Energy Source for Developing Countries: Opportunities and Challenges. *Renew. Sustain. Energy Rev.* **2014**, *31*, 846–859. <https://doi.org/10.1016/j.rser.2013.12.015>.
11. Abanades, S.; Abbaspour, H.; Ahmadi, A.; Das, B.; Ehyaei, M.A.; Esmaeilion, F.; El Haj Assad, M.; Hajilounezhad, T.; Jamali, D.H.; Hmida, A.; et al. A Critical Review of Biogas Production and Usage with Legislations Framework across the Globe. *Int. J. Environ. Sci. Technol.* **2022**, *19*, 3377–3400. <https://doi.org/10.1007/s13762-021-03301-6>.
12. Kemausuor, F.; Adaramola, M.S.; Morken, J. A Review of Commercial Biogas Systems and Lessons for Africa. *Energies* **2018**, *11*, 2984. <https://doi.org/10.3390/en11112984>.
13. Zaken, M. van A. Africa Biogas Partnership Programme (ABPP) Phase 2–Effect Evaluation–Report–Government.NL. Available online: <https://www.government.nl/documents/reports/2019/05/13/africa-biogas-partnership-programme-abpp-phase-2---effect-evaluation> (accessed on 22 April 2026).
14. Bundhoo, Z.M.A.; Surroop, D. Evaluation of the Potential of Bio-Methane Production from Field-Based Crop Residues in Africa. *Renew. Sustain. Energy Rev.* **2019**, *115*, 109357. <https://doi.org/10.1016/j.rser.2019.109357>.
15. Roopnarain, A.; Adeleke, R. Current Status, Hurdles and Future Prospects of Biogas Digestion Technology in Africa. *Renew. Sustain. Energy Rev.* **2017**, *67*, 1162–1179. <https://doi.org/10.1016/j.rser.2016.09.087>.
16. Rupf, G.V.; Bahri, P.A.; de Boer, K.; McHenry, M.P. Barriers and Opportunities of Biogas Dissemination in Sub-Saharan Africa and Lessons Learned from Rwanda, Tanzania, China, India, and Nepal. *Renew. Sustain. Energy Rev.* **2015**, *52*, 468–476. <https://doi.org/10.1016/j.rser.2015.07.107>.
17. Balgah, R.A.; Ketuama, C.T.; Ngwabie, M.N.; Roubik, H. Africa's Energy Availability-Deficiency Paradox: Lessons from Small-Scale Biogas Technology and Policy Implications. *Environ. Dev. Sustain.* **2024**, *26*, 29695–29713. <https://doi.org/10.1007/s10668-023-03810-z>.

18. Brahmi, M.; Bruno, B.; Dhayal, K.S.; Esposito, L.; Parziale, A. From Manure to Megawatts: Navigating the Sustainable Innovation Solution through Biogas Production from Livestock Waste for Harnessing Green Energy for Green Economy. *Heliyon* **2024**, *10*, e34504. <https://doi.org/10.1016/j.heliyon.2024.e34504>.
19. Ghimire, P.C. SNV Supported Domestic Biogas Programmes in Asia and Africa. *Renew. Energy* **2013**, *49*, 90–94. <https://doi.org/10.1016/j.renene.2012.01.058>.
20. Ketuama, C.T.; Mazancová, J.; Roubík, H. Impact of Market Constraints on the Development of Small-Scale Biogas Technology in Sub-Saharan Africa: A Systematic Review. *Environ. Sci. Pollut. Res.* **2022**, *29*, 65978–65992. <https://doi.org/10.1007/s11356-022-22262-y>.
21. International Energy Agency Outlook for Biogas and Biomethane: Prospects for Organic Growth. Available online: https://www.oecd.org/en/publications/outlook-for-biogas-and-biomethane_040c8cd2-en.html (accessed on 27 April 2026).
22. Kelebe, H.E.; Ayimut, K.M.; Berhe, G.H.; Hints, K. Determinants for Adoption Decision of Small Scale Biogas Technology by Rural Households in Tigray, Ethiopia. *Energy Econ.* **2017**, *66*, 272–278. <https://doi.org/10.1016/j.eneco.2017.06.022>.
23. Berhe, M.; Hoag, D.; Tesfay, G.; Keske, C. Factors Influencing the Adoption of Biogas Digesters in Rural Ethiopia. *Energy Sustain. Soc.* **2017**, *7*, 10. <https://doi.org/10.1186/s13705-017-0112-5>.
24. Kamp, L.M.; Bermúdez Forn, E. Ethiopia's Emerging Domestic Biogas Sector: Current Status, Bottlenecks and Drivers. *Renew. Sustain. Energy Rev.* **2016**, *60*, 475–488. <https://doi.org/10.1016/j.rser.2016.01.068>.
25. Şenol, H.; Çolak, E.; Oda, V. Forecasting of Biogas Potential Using Artificial Neural Networks and Time Series Models for Türkiye to 2035. *Energy* **2024**, *302*, 131949. <https://doi.org/10.1016/j.energy.2024.131949>.
26. Yan, B.; Yan, J.; Li, Y.; Qin, Y.; Yang, L. Spatial Distribution of Biogas Potential, Utilization Ratio and Development Potential of Biogas from Agricultural Waste in China. *J. Clean. Prod.* **2021**, *292*, 126077. <https://doi.org/10.1016/j.jclepro.2021.126077>.
27. Pence, I.; Kumaş, K.; Cesmeli, M.S.; Akyüz, A. Future Prediction of Biogas Potential and CH₄ Emission with Boosting Algorithms: The Case of Cattle, Small Ruminant, and Poultry Manure from Turkey. *Environ. Sci. Pollut. Res.* **2024**, *31*, 24461–24479. <https://doi.org/10.1007/s11356-024-32666-7>.
28. Mukasine, A.; Sibomana, L.; Jayavel, K.; Nkurikiyeyezu, K.; Hitimana, E. Maximizing Biogas Yield Using an Optimized Stacking Ensemble Machine Learning Approach. *Energies* **2024**, *17*, 364. <https://doi.org/10.3390/en17020364>.
29. Pala, Z. Data-Centric Evaluation of Machine Learning for Biogas Production Prediction. *Int. J. Energy Res.* **2026**, *2026*, 9261434. <https://doi.org/10.1155/er/9261434>.
30. IRENA Measurement and Estimation of Off-Grid Solar, Hydro and Biogas Energy. Available online: <https://www.irena.org/Publications/2018/Dec/Measurement-and-estimation-of-off-grid-solar-hydro-and-biogas-energy> (accessed on 23 May 2026).
31. IRENA Off-Grid Renewable Energy Statistics 2019. Available online: <https://www.irena.org/Publications/2019/Dec/Off-grid-renewable-energy-statistics-2019> (accessed on 23 May 2026).
32. IRENA Off-Grid Renewable Energy Statistics 2020. Available online: <https://www.irena.org/Publications/2020/Dec/Off-grid-renewable-energy-statistics-2020> (accessed on 23 May 2026).
33. IRENA Off-Grid Renewable Energy Statistics 2021. Available online: <https://www.irena.org/Publications/2021/Dec/Off-grid-renewable-energy-statistics-2021> (accessed on 23 May 2026).
34. IRENA Off-Grid Renewable Energy Statistics 2022. Available online: <https://www.irena.org/Publications/2022/Dec/Off-grid-renewable-energy-statistics-2022> (accessed on 23 May 2026).
35. IRENA Off-Grid Renewable Energy Statistics 2023. Available online: <https://www.irena.org/Publications/2023/Dec/Off-grid-Renewable-Energy-Statistics-2023> (accessed on 23 May 2026).
36. IRENA Off-Grid Renewable Energy Statistics 2024. Available online: <https://www.irena.org/Publications/2024/Dec/Off-grid-Renewable-Energy-Statistics-2024> (accessed on 23 May 2026).
37. IRENA Off-Grid Renewable Energy Statistics 2025. Available online: <https://www.irena.org/Publications/2025/Dec/Off-grid-Renewable-Energy-Statistics-2025> (accessed on 23 May 2026).
38. Food and Agriculture Organization of the United Nations FAOSTAT Database. Available online: <https://www.fao.org/faostat/en/#data> (accessed on 23 May 2026).
39. U.S. Department of Agriculture, Economic Research Service International Agricultural Productivity Data Product. Available online: <https://www.ers.usda.gov/data-products/international-agricultural-productivity> (accessed on 23 May 2026).
40. The World Bank Group World Bank Open Data. Available online: <https://data.worldbank.org/> (accessed on 23 May 2026).

41. Tolessa, A.; Bélières, J.-F.; Salgado, P.; Raharimalala, S.; Louw, T.M.; Goosen, N.J. Assessment of Agricultural Biomass Residues for Anaerobic Digestion in Rural Vakinankaratra Region of Madagascar. *Bioenergy Res.* **2022**, *15*, 1251–1264. <https://doi.org/10.1007/s12155-021-10336-7>.
42. Tolessa, A. Bioenergy Potential from Crop Residue Biomass Resources in Ethiopia. *Heliyon* **2023**, *9*, e13572. <https://doi.org/10.1016/j.heliyon.2023.e13572>.
43. Mboumboue, E.; Njomo, D. Biomass Resources Assessment and Bioenergy Generation for a Clean and Sustainable Development in Cameroon. *Biomass Bioenerg.* **2018**, *118*, 16–23. <https://doi.org/10.1016/j.biombioe.2018.08.002>.
44. Gutu Sakketa, T. Urbanisation and Rural Development in Sub-Saharan Africa: A Review of Pathways and Impacts. *Res. Glob.* **2023**, *6*, 100133. <https://doi.org/10.1016/j.resglo.2023.100133>.
45. Ndambi, O.A.; Pelster, D.E.; Owino, J.O.; de Buissonjé, F.; Vellinga, T. Manure Management Practices and Policies in Sub-Saharan Africa: Implications on Manure Quality as a Fertilizer. *Front. Sustain. Food Syst.* **2019**, *3*, 29. <https://doi.org/10.3389/fsufs.2019.00029>.
46. Clemens, H.; Bailis, R.; Nyambane, A.; Ndung'u, V. Africa Biogas Partnership Program: A Review of Clean Cooking Implementation through Market Development in East Africa. *Energy Sustain. Dev.* **2018**, *46*, 23–31. <https://doi.org/10.1016/j.esd.2018.05.012>.
47. Bond, T.; Templeton, M.R. History and Future of Domestic Biogas Plants in the Developing World. *Energy Sustain. Dev.* **2011**, *15*, 347–354. <https://doi.org/10.1016/j.esd.2011.09.003>.
48. Mutahi, P.M.; Kimutai, S.K.; Adaramola, M.S. Barriers to the Growth and Implementation of Biogas Technology: The Kenyan Experience. *Energy Sci. Eng.* **2025**, *13*, 5268–5283. <https://doi.org/10.1002/ese3.70237>.
49. Getaneh, A.; Eba, K.; Tucho, G.T. Assessment of Biomass Energy Potential for Biogas Technology Adoption and Its Determinant Factors in Rural District of Limmu Kossa, Jimma, Ethiopia. *Energies* **2024**, *17*, 2176. <https://doi.org/10.3390/en17092176>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.