Global Biogeochemical Cycles

REVIEW ARTICLE

10.1029/2026GB009141

# Nitrous Oxide Emissions Across Sub-Saharan Africa: Meta-Analysis and Data-Driven Modeling

P. Agredazywczuk[1,2], T. Ouma[1,2,3], M. Barthel[3], A. Otinga[4], R. Njoroge[4], K. Butterbach-Bahl[5,6], R. Daelman[7,8], J. E. Hickman[9,10], W. Ibrahim[11], M. Laub[3], S. Leitner[11], A. Shumba[12,13,14], K. L. Tully[15], S. Wachiye[16,17], J. Zheng[18], M. Bauters[8], R. Kiese[5], R. Cardinael[13,14,19], C. Andres[3], R. Balokah[20], K. S. Lourenço[21], K. L. Ouattara[22], J. Six[3], and E. Harris[2]

[1]Swiss Data Science Center (EPFL and ETH Zurich), Zurich, Switzerland, [2]Climate and Environmental Physics, University of Bern, Bern, Switzerland, [3]Department of Environmental Systems Science, Sustainable Agroecosystems Group, ETH Zurich, Switzerland, [4]Department of Soil Science, University of Eldoret, Eldoret, Kenya, [5]Institute of Meteorology and Climate Research, Atmospheric Environmental Research (IMK-IFU) Karlsruhe Institute of Technology (KIT), Garmisch-Partenkirchen, Germany, [6]Pioneer Center Land-CRAFT, Department of Agroecology, Aarhus University, Aarhus, Denmark, [7]Department of Green Chemistry and Technology, ISOFYS, Ghent University, Ghent, Belgium, [8]Q-Forest lab, Department of Environment, Ghent University, Ghent, Belgium, [9]Center for Climate Systems Research, Columbia University, New York, NY, USA, [10]NASA Goddard Institute for Space Studies, New York, NY, USA, [11]Mazingira Centre for Environmental Research and Education, International Livestock Research Institute (ILRI), Nairobi, Kenya, [12]Fertilizer, Department of Research and Specialist Services, Farm Feeds and Remedies Institute, Ministry of Lands, Agriculture, Harare, Zimbabwe, [13]CIRAD, UPR AIDA, Harare, Zimbabwe, [14]Department of Plant Production Sciences and Technologies, University of Zimbabwe, Harare, Zimbabwe, [15]University of Maryland: College Park, College Park, MD, USA, [16]Department of Agro-Forestry, Environmental Studies, and Integrated Natural Resources, University of Kabianga, Kericho, Kenya, [17]Department of Agricultural Sciences, University of Helsinki, Helsinki, Finland, [18]Crop, Livestock and Environment Division, Japan International Research Center for Agricultural Sciences, Tsukuba, Japan, [19]AIDA, University of Montpellier, CIRAD, Montpellier, France, [20]Agroforestry Department, Faculty of Renewable Natural Resources (FRNR), Kwame Nkrumah University of Science and Technology, Kumasi, Ghana, [21]Plant Production Systems, Wageningen University & Research, Wageningen, The Netherlands, [22]Nestle Institute of Agricultural Sciences, Zambakro, Switzerland

**Key Points:**

- Machine learning shows climate change (SSP5-8.5) may cut Sub-Saharan Africa (SSA) $N_2O$ emissions by 4%–37% due to drier soils inhibiting microbial production
- Natural ecosystems may emit less $N_2O$ under severe climate change (SSP5-8.5), but tripling fertilizer use could increase cropland emissions by 6%–23%
- SSA contributes 253–538 Gg N $yr^{-1}$ (1%–3% global), but improved predictions require urgent investment in monitoring infrastructure

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

P. Agredazywczuk,
phillip.agredazywczuk@unibe.ch

**Citation:**

Agredazywczuk, P., Ouma, T., Barthel, M., Otinga, A., Njoroge, R., Butterbach-Bahl, K., et al. (2026). Nitrous oxide emissions across Sub-Saharan Africa: Meta-analysis and data-driven modeling. *Global Biogeochemical Cycles*, *40*, e2026GB009141. https://doi.org/10.1029/2026GB009141

Received 20 FEB 2026
Accepted 21 MAY 2026

**Author Contributions:**

**Conceptualization:** E. Harris
**Data curation:** P. Agredazywczuk, T. Ouma, A. Otinga, R. Njoroge, K. Butterbach-Bahl, R. Daelman, J. E. Hickman, W. Ibrahim, M. Laub, A. Shumba, K. L. Tully, S. Wachiye, J. Zheng, M. Bauters, R. Kiese, R. Cardinael, C. Andres, R. Balokah, K. S. Lourenço, K. L. Ouattara, J. Six, E. Harris
**Formal analysis:** P. Agredazywczuk
**Investigation:** P. Agredazywczuk
**Methodology:** P. Agredazywczuk, E. Harris



## Abstract

Food security and avoiding land use change in Sub-Saharan Africa (SSA) requires increasing agricultural productivity, necessitating greater fertilizer use. This may increase soil nitrous oxide ($N_2O$) emissions, a potent greenhouse gas. This study used Machine learning (ML) models to predict $N_2O$ emissions under future climatic and fertilizer scenarios across SSA. Three models were trained (Random Forest (RF), XGBoost (XGB), and feedforward neural networks (FNN)) on existing forest, grassland, and cropland $N_2O$ measurements. The analysis identified the main drivers influencing $N_2O$ emissions: temperature, soil moisture, rainfall, and fertilizer (cropland). SSA $N_2O$ emissions totaled 253–538 Gg N $yr^{-1}$ (1 Gg = $10^9$ g), with forests contributing 119–342 Gg N $yr^{-1}$, grasslands 70–132 Gg N $yr^{-1}$, and croplands 63–64 Gg N $yr^{-1}$. Ranges reflect the full uncertainty across all models. Unexpectedly, severe climate change (SSP5-8.5 scenario) may decrease total $N_2O$ emissions by 4%–37% across SSA, possibly due to drier soils. However, when climate change was combined with tripled fertilizer use (0–63 to 0–189 kg N $ha^{-1}$ $yr^{-1}$), the models predicted a wide range of cropland emission increases of 6%–23% (XGB–RF), with FNN predicting 139% from baseline projections. These findings highlight that while climate change may reduce overall $N_2O$ emissions from forests and grasslands, agricultural intensification will likely become an increasingly significant emission source. To improve prediction accuracy, more comprehensive $N_2O$ monitoring across SSA is needed. This work underscores the need for international investment in monitoring infrastructure and data repositories to guide sustainable agricultural development and SSA climate policy.

## Plain Language Summary

Sub-Saharan Africa (SSA) faces a critical challenge: achieving food security while protecting the environment. Farmers in SSA will need to increase fertilizer use to grow enough food, but this could increase nitrous oxide ($N_2O$) emissions, a strong greenhouse gas. To understand this trade-off, we used three ML models: Random Forest (RF), XGBoost (XGB), and feedforward neural networks (FNN) to predict future $N_2O$ emissions across SSA under worst-case climate and triple fertilizer scenarios. We trained models on measurements from forests, grasslands, and croplands, finding that temperature, soil moisture, rainfall, and fertilizer use were key factors controlling $N_2O$ production. Currently, SSA releases about 253,000–538,000 tonnes of nitrogen as $N_2O$ annually. Most comes from natural forests (47%–64%) and

**Supervision:** J. Six, E. Harris
**Validation:** P. Agredazywczuk
**Visualization:** P. Agredazywczuk
**Writing – original draft:** P. Agredazywczuk
**Writing – review & editing:** A. Otinga, R. Njoroge, K. Butterbach-Bahl, R. Daelman, J. E. Hickman, W. Ibrahim, M. Laub, A. Shumba, K. L. Tully, S. Wachiye, J. Zheng, M. Bauters, R. Kiese, R. Cardinael, K. S. Lourenço, J. Six, E. Harris

grasslands (25%–28%), not croplands (12%–25%). Models predict that severe climate change alone might reduce total emissions by 4%–37%, likely because soils will become drier. However, if farmers triple their fertilizer use, cropland emissions could increase by 6%–23% (XGB–RF), with FNN predicting 139% from baseline projections. This suggests that while climate change may reduce "natural" $N_2O$ emissions, agricultural intensification will become increasingly important. International investment in monitoring infrastructure could help SSA pursue sustainable agricultural development, balancing food security with climate protection.

## 1. Introduction

Sub-Saharan Africa (SSA) is the region most affected by food insecurity globally, with approximately 30% of the population facing food shortages due to persistently low crop yields (FAO, 2018). Crop production in SSA is dominated by smallholder farms, defined as those covering an area <2 ha (Lowder et al., 2016). In some SSA countries, smallholder farms contribute up to 90% of agricultural production (Wiggins, 2009), making them crucial to national-level food security and global agricultural product supply, such as coffee, tea, and cocoa. Important crops for food production include cereals (maize and sorghum), legumes (cowpea), and tubers (potatoes and cassava) (Chemura et al., 2024; van Ittersum et al., 2016). Among these, maize is one of the most dominant crops, grown across over 40 million hectares (approximately 18% of total arable land) and serving as the primary cereal crop in over half of SSA countries (FAO, 2021). Yields from smallholder farms in SSA are generally low due to low mechanization, limited irrigation, low pest control measures, and low fertilizer use (Mueller et al., 2012; Sheahan & Barrett, 2017; Van Koppen, 2003). Fertilizer application rates in SSA are estimated to average only 16 kg N $ha^{-1}$ $yr^{-1}$, compared to over 100 kg N $ha^{-1}$ $yr^{-1}$ in Europe and North America, and over 150 kg N $ha^{-1}$ $yr^{-1}$ in China (Rurinda et al., 2020). To meet its regional food security needs, SSA must either increase its agricultural output or rely more heavily on food imports (van Ittersum et al., 2016). Moreover, with Africa's population projected to increase from about 1.3 billion in 2020 to around 2.5–2.7 billion by 2050, expanding SSA's agricultural productivity will be crucial for meeting rapidly growing regional demands for food and animal feed (Godfray et al., 2010; United Nations, 2019). Expanding agricultural land area at the expense of natural habitat, such as forests and grasslands, poses a significant threat to the continent's rich biodiversity (Jayne & Sanchez, 2021; van Ittersum et al., 2016). Therefore, intensifying agricultural production in SSA will undoubtedly rely on increased fertilizer use, alongside other sustainable intensification measures (Falconnier et al., 2023).

The impact of increased fertilizer use on nitrous oxide ($N_2O$) emissions from tropical soils remains uncertain. However, previous estimates indicated that agricultural $N_2O$ emissions accounted for 42% of total $N_2O$ emissions in Africa (Hickman et al., 2011). $N_2O$ is a strong greenhouse gas (GHG), with a global warming potential 273 times greater than carbon dioxide ($CO_2$) over 100 years (IPCC, 2023), and it is also a significant ozone-depleting substance (Portmann et al., 2012). There have been limited field-based estimates of $N_2O$ from SSA croplands, especially under varying fertilisation regimes or at high temporal resolution (Cardinael et al., 2024; Hickman et al., 2014, 2015; Ntinyari et al., 2023; Tully et al., 2023; Zheng et al., 2019). Fertilizer inputs above crop demand have been shown to have a non-linear effect on direct $N_2O$ soil emissions globally and in SSA (Hickman et al., 2015; Shcherbak et al., 2014; Van Groenigen et al., 2010). Kim et al. (2013) showed a sigmoidal response of $N_2O$ emissions to increasing N fertilizer inputs, suggesting that inputs >150 kg N $ha^{-1}$ may lead to a disproportionate increase in $N_2O$ emissions in SSA. Additionally, field-based $N_2O$ estimates from natural ecosystems are scarce (Kim et al., 2016). The lack of field-based estimates of $N_2O$ introduces uncertainties in calculating current and future $N_2O$ emissions in SSA, limiting our ability to predict the effects of agricultural intensification or climate change, and thus preventing the development of effective mitigation strategies.

$N_2O$ production in soils is primarily driven by two key biogeochemical processes: nitrification and denitrification (Braker & Conrad, 2011; Syakila & Kroeze, 2011). These nitrogen (N) cycling processes are regulated by several parameters, including soil moisture, which influences oxygen availability (Davidson et al., 2000), temperature (Rosenstock et al., 2016), pH (Baggs et al., 2010), organic matter (Zhuang et al., 2012), soil type (Pelster et al., 2012; Zhu et al., 2020), and plant processes (Basche et al., 2014). The interaction of N inputs, microbial and plant processes, and soil characteristics lead to considerable variability in $N_2O$ production across space and time, often resulting in "hotspots" or "hot moments" (Butterbach-Bahl et al., 2013; Groffman et al., 2009). Capturing these emission patterns requires high-resolution instruments capable of detecting spatial heterogeneity and temporal dynamics, as $N_2O$ emissions can vary by orders of magnitude within meters (Kravchenko & Robertson, 2015) and over short time scales, such as days or even hours (Barton et al., 2015). Moreover, peaks in $N_2O$ fluxes can show a lag time between environmental triggers and emission responses, days after peaks in soil moisture or significant precipitation events (Rautakoski et al., 2024). However, $N_2O$ measurements in SSA are scarce, with limited measurement infrastructure in the region and low temporal resolution measurements sparse, making it difficult to constrain $N_2O$ drivers and emissions. Consequently, uncertainties in $N_2O$ processes, drivers, and total emissions remain high in the region (Tian et al., 2024).

Modelling approaches can fill the gaps caused by measurement challenges, particularly in data-limited regions. Most SSA countries rely on IPCC (2006) Tier 1 emission factors (EFs) for $N_2O$ emission estimates, which derive primarily from temperate ecosystem studies (Bouwman et al., 2002). However, these EFs often prove inadequate for (sub-)tropical regions due to distinct environmental parameters, including soil carbon (C) content, texture, and pH that differ significantly from temperate zones (Charles et al., 2017; Rochette et al., 2018; Shcherbak et al., 2014). In contrast, process-based models can estimate $N_2O$ emissions by simulating soil biogeochemical activity, primarily C and N cycling, while considering soil characteristics. Several such models are available, including DNDC (Li et al., 1992), DayCENT (Grosso et al., 2006), and MITERRA-EUROPE (Velthof et al., 2007). However, these models were primarily developed based on data from temperate zones, particularly European and North American grasslands, limiting their applicability to (sub-)tropical soils and crops (Costa et al., 2021; Laub et al., 2024; Levy et al., 2007; Vuichard et al., 2007). (Sub-)tropical soils are highly weathered and possess distinct characteristics compared to temperate soils, including differences in mineralogical (Hodnett & Tomasella, 2002; Six et al., 2002), hydrological and biogeochemical properties (Tomasella & Hodnett, 2004; Wohl et al., 2012), and acidic soil pH (Vitousek, 1984; Xu et al., 2013). These differences can influence nitrification and denitrification rates and consequently $N_2O$ emissions (Xu & Cai, 2007; Xu et al., 2013). Moreover, these models require extensive calibration data for reliable results, creating additional uncertainty in regions where measurements are limited (Berardi et al., 2020; Ehrhardt et al., 2018; Gaillard et al., 2018; Laub et al., 2024). This mismatch between temperate-based process models and (sub-)tropical conditions creates substantial uncertainty in regional emission estimates.

Machine learning (ML) models offer a data-driven alternative by learning patterns and relationships from observations (Aderele et al., 2025). Recent studies have demonstrated the effectiveness of ML models in predicting agricultural $N_2O$ emissions. These include Random Forest (RF; Saha et al., 2021), which is an ensemble method that combines multiple decision tress to capture non-linear relationships and is robust to overfitting; extreme gradient boosting (XGBoost; Kim et al., 2021), which is an ensemble method that sequentially corrects prediction errors to try and achieve high accuracy; and deep neural networks (DNN; Hamrani et al., 2020), which learn hierarchical data representations to capture non-linear patterns. While these approaches show promise, their training and testing have been limited to intensively monitored agricultural sites with high-resolution data. The scarcity of such data sets in SSA presents both a challenge and an opportunity to evaluate the predictive capabilities of ML models under data-limited conditions and other environments.

This study will analyse available literature and gather temporal $N_2O$ data across SSA environments, including different N fertilizer treatments (0–200 kg N $ha^{-1}$) for diverse land uses such as natural (forests and grasslands) or agricultural sites, particularly croplands. The aim is to identify $N_2O$ emission drivers and to optimise and compare three ML methods (RF, XGBoost, and feedforward neural network (FNN)) for predictive modelling and spatial upscaling. The specific study objectives are to: (a) assess ML model performance; (b) identify key drivers of $N_2O$ emissions across environments using ML modelling; (c) quantify the annual $N_2O$ emissions; and (d) model regional $N_2O$ production scenarios under different fertilizer and climate scenarios.

## 2. Methodology

### 2.1. Data Collection and Interpolation

$N_2O$ emission data were retrieved from peer-reviewed literature by searching Scopus, Google Scholar and Web of Science published between 2000 and 2023, using keywords such as “SSA,” ‘$N_2O$′ or “nitrous oxide” and various soil environment categories like “croplands, “savanna,” “forest,” “plantations,” or “grazing land.” Additional

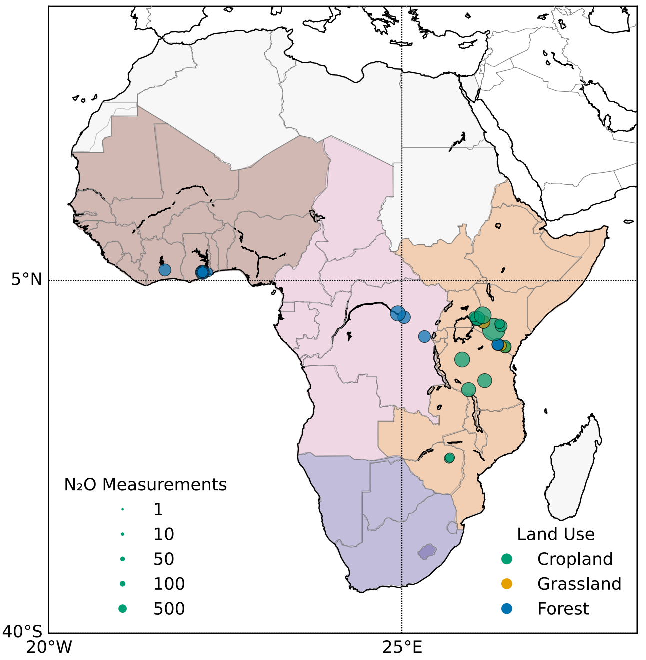


**Figure 1.** Study site locations and relative number of $N_2O$ measurements. The color shades represent Sub-Saharan Africa regions; Brown—Western Africa, Pink—Central Africa, Orange—Eastern Africa, Purple—Southern Africa. '$N_2O$ Measurements' refers to the total number of measurements at a site, including different treatments. Data collected across 14 cropland sites, 8 forest sites, and 4 grassland sites. Some studies investigated multiple land use types.

publications were retrieved from databases compiled by Leitner et al. (2020). The selection criteria for studies included the following: only field-based studies were considered, excluding laboratory-based studies or modeled results; $N_2O$ measurements were conducted for at least 3 months; and fertilisation regimes were reported in croplands while intercropping, conservation agriculture and biomass residue were considered beyond the scope of this study and excluded. A total of 45 peer-reviewed publications were identified across the different environment categories. Most cropland studies focussed on maize crops, a staple crop in SSA; however, some studies did contain different crop types such as potatoes. As most studies reported aggregated measurements, corresponding authors of the retrieved studies were contacted to provide continuous or semi-continuous $N_2O$ measurements and any continuous or semi-continuous meteorological or soil environmental microsite measurements collected during the study, receiving 16 "raw" published or unpublished data sets from the authors contacted from the 45 publications (Figure 1; Barthel et al. (2022); Daelman et al. (2024); Gütlein et al. (2018); Hickman et al. (2014), Hickman et al. (2015); Laub et al. (2023); Shumba et al. (2023); Wanyama et al. (2018); Zheng et al. (2019); Wachiye et al. (2020). Within these 16 data sets, 77 data collections were available, which refer to different subsites or treatments within a study site. For example, if a cropland study had four different fertilizer treatments, each would be considered a data collection. Each data collection contained multiple $N_2O$ flux measurements over time. The total number of individual observations available for forests, grasslands, and croplands were 1,584, 225, and 3,574, respectively.

The additional data provided varied between studies and was often sparsely reported. Therefore, soil and meteorological variables in the SSA datasets were interpolated using ERA5-Land Reanalysis (Soci et al., 2024), which provides high-resolution data at a regular latitude-longitude grid (0.25° × 0.25°) at 9 km resolution and daily temporal resolution. For study sites where geographical coordinates were not reported in the original literature, coordinates were georeferenced using Google maps (http://www.google.com/maps). Soil properties such as pH, texture, or soil organic carbon (SOC) were not considered as it was not the focus of the study, but it is acknowledged that these are important variables to consider for $N_2O$ emission production.

The ML models were trained using 20 variables, which are shown in Table S2 in Supporting Information S1, including the ERA5 soil and climate variables, site-specific temporal variables (days since rainfall and fertilisation), and a fertilizer decay function (Equation 1). Overall comparisons of ERA5 data and available SSA variable observations are provided in Table S3 in Supporting Information S1. The fertilizer-related variables were only considered for croplands. Using histogram analysis on the ERA5 rainfall, rainfall events were defined with a threshold of >0.0042 mm. Fertilisation dates were collected from the published literature and direct author correspondence took place for unpublished data sets. Nutrient dynamics were modeled using an exponential fertilizer decay function, implementing an annual (365-day) half-life and accounting for two fertilisation dates, commonly applied in the cropland studies used in this study.

$$F(t) = F_0 \cdot e^{-\frac{\ln(2)\cdot t}{t_{1/2}}} \tag{1}$$

Where $F(t)$ represents the fertilizer application at time $t$ (days), $F_0$ is the initial fertilizer concentration following application, $e$ is Euler's number (2.71828), $t_{1/2}$ denotes the half-life of the fertilizer in the soil matrix (365 days in our system), and $t$ is the elapsed time since fertilizer application (days).

Seasonality was incorporated into modeling using sine and cosine transformations of the day/month of the year (values between 0 and 2π—mapping one complete year to one full sine/cosine cycle), enabling the ML methods to recognize cyclical data patterns. Lastly, the study sites were classified into three land use types (forest, grassland, and cropland), and separate ML models were trained for each type. For a full-flow diagram of the methodology, refer to Figure S1 in Supporting Information S1.

### 2.2. Model Descriptions

ML methods were used to model $N_2O$ emissions from the SSA sites between 2005 and 2025. A temporal block-based split was implemented to divide the data sets into training (80%) and test (20%) sets for each land use type. This approach ensured temporal continuity within the blocks while preventing data leakage from adjacent observations. Block sizes were determined based on the total data set size for each land use, the number of blocks, and the minimum block size (ranging from 2 to 10 $N_2O$ flux observations) were calculated to achieve the 80/20 split. Block sizes were then varied around the mean using controlled randomness (±10%) to prevent uniform sampling. Blocks were distributed throughout the time series sequentially with gaps to ensure the training and test sets were not temporally adjacent to each other, preventing autocorrelation. A single test-train split was created for each land use, with the following number of individual $N_2O$ daily flux observations: forests ($N = 1584$ observations from 8 study sites and 17 data collections), grasslands ($N = 225$ observations from 4 study sites and 5 data collections), and croplands ($N = 3574$ observations from 14 study sites and 55 data collections). For forests, 162 blocks were created; for grasslands, 22 blocks; and for croplands, 357 blocks (Figure S2 in Supporting Information S1). While multiple data collections from the same study site may share similar environmental conditions, the temporal block-based approach minimizes overfitting by ensuring test blocks are temporally separated from training data.

To note, correlation analysis identified correlations between air and soil temperatures (and amongst soil temperatures) at different depths ($r = 0.86$–$0.95$), and amongst soil moisture depths ($r = 0.65$–$0.74$). Surprisingly, rainfall and soil moisture were not highly correlated ($r = 0.013$–$0.073$). Instead of removing correlated variables before model training, they were retained for several reasons. Tree-based models like RF and XGB can effectively handle correlated predictors, selecting the most important variable at each decision node, and reducing the importance of redundant variables in subsequent splits. Moreover, including multiple correlated measurements can improve predictive performance by capturing small variations and reducing overfitting through variable diversity.

#### 2.2.1. Random Forest (RF)

RF combines multiple decision trees through bagging (bootstrap aggregation), building independent trees using random subsets of data and features, and aggregating predictions to produce the final output. RF is known for its robust performance with limited data, variable importance ranking, computational efficiency, and the ability to capture complex non-linear interactions between variables (Khalil et al., 2023; Philibert et al., 2013; Saha et al., 2021). Feature selection was performed using recursive feature elimination with cross-validation to identify optimal predictor variables. The top eight features were identified based on importance scores, and then feature combinations were evaluated using 5-fold cross-validation. Feature combinations were assessed using several metrics (Section 2.3), with final feature importance scores weighted between cross-validation (30%) and validation set performance (70%). Hyperparameter tuning was conducted using a grid search approach with early stopping criteria, including number of trees (50–200), maximum tree depth (3–10), and minimum samples for splits (2–10). The Python package `sklearn` was used.

#### 2.2.2. XGBoost (XGB)

XGB is a boosting tree model algorithm which uses regression trees based on the classification and regression tree framework. It creates a scalable gradient tree ensemble model that trains weak learners sequentially to optimize model performance (Chen & Guestrin, 2016; Wan et al., 2023). Following the same temporal block-based split and feature selection methodology as described for RF, XGB hyperparameter tuning was conducted using a grid search approach with early stopping criteria. The hyperparameters optimized included the number of estimators (50–200), maximum tree depth (3–10), learning rate (0.01–0.3), minimum child weight (1–5), subsample ratio (0.8–1.0), and column sampling per tree (0.8–1.0). Early stopping was implemented with a patience of 50 iterations to prevent overfitting. The Python package `xgboost` was used.

#### 2.2.3. Feedforward Neural Network (FNN)

FNN was also implemented to model $N_2O$ emissions. Neural networks are attractive models due to their hierarchical architecture, composed of multiple interconnected layers with numerous neurons, enabling them to learn and capture complex non-linear relationships in data. FNNs use unidirectional information flow (no feedback loops) and non-linear activation functions, such as ReLU (Rectified Linear Unit) or sigmoid, at each neuron. These functions allow the network to model complex non-linear patterns in data through simpler functions across multiple layers (Stursa & Dolezel, 2019; Wang et al., 2019).

The FNN architecture consisted of two hidden layers with ReLU activation functions, batch normalization, and dropout layers (rate: 0.1–0.2) to prevent overfitting. The model was optimized using the Adam optimizer with learning rates between 0.001 and 0.01, and hidden layer sizes of 32 or 64 nodes. Adam optimizer is an algorithm that helps the neural network learn better by adjusting the learning rate for each parameter during training (Kingma & Ba, 2014). Model training incorporated early stopping with patience of 50 epochs to prevent overfitting, and gradients were clipped to ensure stable training. The input features were standardised using z-score normalization, and the target variable was log-transformed ($\log_1(x + 1)$) to handle the skewed distribution of $N_2O$ emissions. Model performance was evaluated using several metrics (Section 2.3) on both validation (20% of training data) and test sets. The Python package `keras` was used.

### 2.3. Predictive Performance Assessment

Final model performance was evaluated using multiple metrics on the training and test sets. The training sets were used to fit the models and assess their ability to learn data patterns, while the independent test set evaluated their generalization ability to unseen data. Model performances were quantified using the coefficient of determination ($R^2$; Equation 2), which measures the proportion of variance in observed emissions explained by the model and indicates overall predictive capability. Mean error (ME; Equation 3) quantifies systematic bias in predictions, showing whether models over– or under-predict emissions. Root mean square error (RMSE; Equation 4) and mean absolute error (MAE; Equation 5) measure prediction accuracy. RMSE is more sensitive to large errors and MAE is a measure of average prediction error. The distance between indices of simulation and observation (DISO; Equation 6) measures whether the models correctly identify the relative emission magnitudes (distinguishing from low vs high periods), with lower values indicating better agreement.

In the equations below, $y_i$ represents observed $N_2O$ flux values, $\hat{y}_i$ represents predicted values, $\bar{y}$ is the mean of observed values, and $n$ is the number of observations. For DISO, $\text{rank}(y_i)$ and $\text{rank}(\hat{y}_i)$ represent the rank positions of observed and predicted values when sorted in ascending order.

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

$$\text{ME} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \tag{3}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{4}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{5}$$

$$\text{DISO} = \frac{1}{n}\sum_{i=1}^{n}\left|\text{rank}(y_i) - \text{rank}(\hat{y}_i)\right| \tag{6}$$

In addition, the models' annual cumulative performance was compared by multiplying the mean of all flux measurements by total annual hours (365 × 24) and then converting from μg $N_2O$-N $m^{-2}$ to kg $N_2O$-N $ha^{-1}$.

For SSA upscaling comparison, agreement among RF, XGB, and FNN predictions was assessed using pixel level inter-model comparisons. Pearson's correlation coefficient ($r$) was used to quantify spatial agreement between model pairs, indicating the extent to which different models captured similar spatial patterns of $N_2O$ emissions.

The root mean square difference (RMSD; Equation 7) was used to quantify the average magnitude of differences between model predictions at the pixel level:

$$\mathrm{RMSD} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i^{(1)} - \hat{y}_i^{(2)}\right)^2} \tag{7}$$

where $\hat{y}_i^{(1)}$ and $\hat{y}_i^{(2)}$ represent predictions from two different models at pixel $i$, and $n$ is the total number of pixels. RMSD reflects both random and systematic differences in predicted emission magnitudes between models.

To facilitate comparison across land uses and scenarios, the normalized RMSD (NRMSD) was calculated by expressing RMSD as a percentage of the mean predicted emission value:

$$\mathrm{NRMSD} = \frac{\mathrm{RMSD}}{\bar{\hat{y}}} \times 100 \tag{8}$$

where $\bar{\hat{y}}$ is the mean of predicted $N_2O$ emissions across all pixels.

Systematic differences between models were quantified using percent bias (PBIAS; Equation 9):

$$\mathrm{PBIAS} = \frac{\sum_{i=1}^{n}\left(\hat{y}_i^{(1)} - \hat{y}_i^{(2)}\right)}{\sum_{i=1}^{n}\hat{y}_i^{(2)}} \times 100 \tag{9}$$

Positive PBIAS values indicate that the first model predicts higher emissions than the second, while negative values indicate underestimation relative to the reference model.

Together, these inter-model metrics enabled quantification of structural uncertainty associated with model selection in the SSA upscaled $N_2O$ emission estimates.

### 2.4. Regional $N_2O$ Emission Upscaling

To upscale $N_2O$ emissions from the individual studies to the full SSA region, a 365-day ”representative” year was created for each grid cell by averaging daily ERA5 reanalysis data over 14 years (2010–2024), corresponding to the temporal range of most $N_2O$ measurements. Annual synthetic N fertilizer application data were obtained from the HaNi dataset (spatial resolution ∼9 km; Tian et al., 2022) and averaged from 2010 to 2019, limited by the temporal coverage, which ceased in 2019. Land cover classification was derived from the 2021 ESA World Cover product (spatial resolution 10 m; Zanaga et al., 2022). The 2021 ESA World Cover focuses on annual croplands and does not distinguish between different crop types. All data sets were regridded to a 2.5 km resolution, applying linear interpolation for the continuous variables (ERA5 and HaNi datasets) and nearest neighbor interpolation for the categorical ESA land cover classification. Using these regridded data sets, the trained ML models were applied to predict daily $N_2O$ emissions for grid cells representing each land use type (forest, cropland, grassland) across SSA in the “representative” year and under the SSP5-8.5 climate and fertilizer scenario. To account for the timing effects of fertilizer application on $N_2O$ emissions, days since fertilizer and fertilizer decay variables were calculated based on the observed application dates in the collected field studies. The observed fertilizer application days were averaged to calculate ’representative fertilizer application days' (day 127 and 160—split application) which were then applied to the ’representative’ year in each HaNi grid cell. The models use these dates to assess fertilizer-timing effects for each grid cell. For SSA pixels with a potential of two cropping seasons, fertilization was assumed to occur only during the main cropping season, as this represents the dominant fertilization pattern in the field studies. The model accounts for interactions between fertilizer and climate and soil variables, allowing $N_2O$ response to fertilizer to vary under different climatic conditions. The SSP5-8.5 climate scenario refers to a future where global social and economic development still relies heavily on fossil fuels, leading to high levels of greenhouse gas emissions and considerable warming (Lee et al., 2021). To simulate the climate scenario, SSA projections from Almazroui et al. (2020) were used, which indicated a boreal winter increase in rainfall and air temperature by 35.2% and 4.1°C, respectively, and a boreal summer increase of 10% and 4.6°C, respectively, by the end of the century (2099).

For implementation purposes, these boreal winter and summer increases were applied to days 305–120 and days 121–304 of the year, respectively. Temperature increases were uniformly applied across all three grid cell temperature variables (temp, tmin, and tmax). Moreover, regression relationships were established between air temperature and soil temperature, and between rainfall and soil moisture, with the resulting regression slopes used to estimate changes in grid cell soil conditions under the SSP5-8.5 climate scenario (Figure S3 in Supporting Information S1). Precipitation patterns were not directly modified in this study, though changes are expected under future climate scenarios. However, precipitation changes over space and time remain uncertain across different climate projections and geographical regions. Grid cell fertilizer application rates were tripled, with the mean baseline application rate at 0.84 kg N $ha^{-1}$ (range: 0–63 kg N $ha^{-1}$) and the mean tripled application rate at 2.5 kg N $ha^{-1}$ (range: 0–189 kg N $ha^{-1}$). To further understand the response of $N_2O$ to N inputs by climatic region and scenario, a fertilizer response ratio was calculated for SSA and SSA regions (Equation 7).

$$\text{Fert. response ratio} = \frac{(\text{climate\_scenario\_triple\_fert} - \text{climate\_scenario})}{(\text{baseline\_triple\_fert} - \text{baseline})} \tag{10}$$

Where climate_scenario_fert3 represents the SSP5-8.5 climate with triple fertilisation scenario; climate_scenario represents the SSP5-8.5 climate only scenario; baseline_fert3 represents the baseline climate with triple fertilisation scenario; and baseline represents the baseline climate only scenario.

## 3. Results

### 3.1. Machine Learning Model Performance

The performances of RF, XGB, and FNN models varied across different land use types in predicting $N_2O$ emissions (Figure 2 and Table 1). For forests, RF and XGB showed similar predictive capability ($R^2 = 0.36$ and 0.40, respectively), while FNN performed considerably worse, with a negative $R^2$ value (−0.20), indicating it failed to capture the variance structure of forest flux observations. For grasslands, the pattern was reversed: FNN showed the strongest predictive capability ($R^2 = 0.55$), followed by RF ($R^2 = 0.45$), while XGB performed the poorest ($R^2 = 0.13$). For croplands, XGB performed best ($R^2 = 0.41$), with RF and FNN showing similar but slightly lower predictive capability ($R^2 = 0.36$ and 0.31, respectively). No single model consistently outperformed the others across all land use types, reflecting the heterogeneity of the underlying data sets rather than an inherent superiority of any one modeling approach.

Considering all performance metrics in Table 1, FNN showed the lowest overall error relative to the magnitude of emissions across all land use types, based on DISO values (1.42–1.71). In contrast, RF and XGB showed substantially higher DISO values for croplands (235.89 and 223.83, respectively) and forests (99.15 and 93.23, respectively), indicating that these models had larger systematic errors in absolute flux magnitudes for these land use types, despite their comparable or superior $R^2$ values. The divergence between $R^2$ and DISO across models highlights that no single metric captures all aspects of model performance, and that model selection involves inherent trade-offs depending on the intended application.

The performance of the models showed a marked contrast between predicting individual $N_2O$ fluxes and annual cumulative emissions across different land use types. For cumulative site emissions (Figure 3), model performance varied considerably by land use. In grasslands, all three models performed well, with FNN achieving the highest predictive capability ($R^2 = 0.78$), followed by RF ($R^2 = 0.67$) and XGB ($R^2 = 0.56$). In croplands, model performance was comparable across approaches, with $R^2$ values ranging from 0.59 to 0.62, indicating similar predictive capability. In contrast, forests showed greater variance in cumulative performance, with only XGB demonstrating reasonable predictive capability ($R^2 = 0.62$), while RF and FNN performed poorly ($R^2 \approx 0$ and −0.10, respectively). These results suggest that cumulative $N_2O$ emissions can be reliably estimated for grassland and cropland systems, whereas estimates for forest sites have greater uncertainty. Despite this, all models struggled to capture short-term variability and peak emission events across land use types (Figure S5 and Figure S6 in Supporting Information S1).

Model selection for upscaling was based on a combined assessment of instantaneous and cumulative flux performance across all land use types and metrics. RF was selected as the primary upscaling model based on its relatively consistent performance across all three land use types for both instantaneous and cumulative fluxes. XGB was retained as a lower-bound sensitivity estimate given its systematically lower predictions, and its

10.1029/2026GB009141

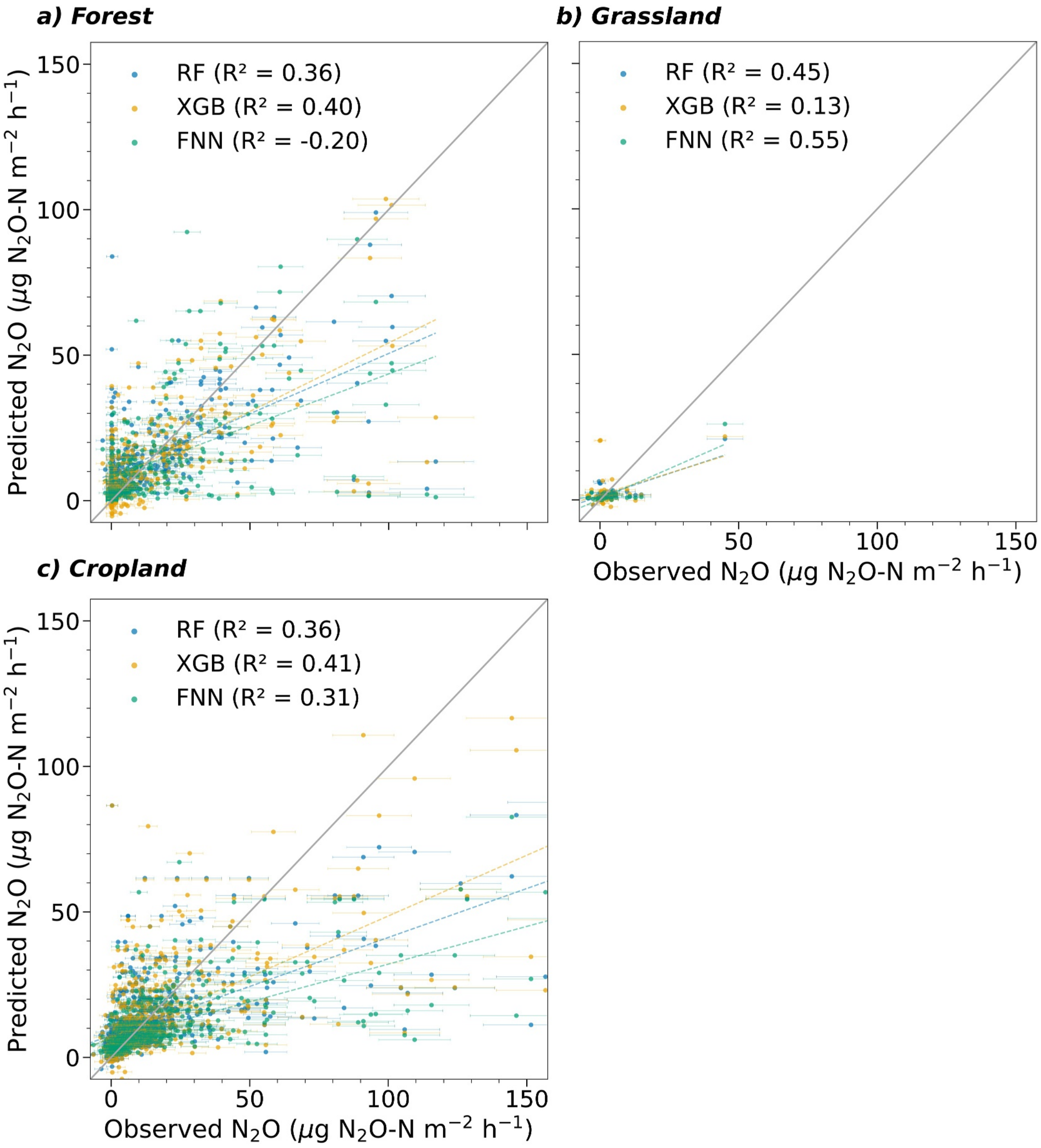


**Figure 2.** Comparison of observed and predicted mean daily $N_2O$ fluxes on the test set: (a) Forest, (b) grassland, and (c) cropland. Error bars represent the standard deviation of observed mean daily $N_2O$ fluxes. For each land use, the machine learning models were trained on all data sets.

reasonable performance for forests and croplands. Although FNN showed strong predictive capability for grasslands and the lowest DISO values overall, its strongly negative $R^2$ for both instantaneous (−0.20) and cumulative (−0.10) forest fluxes indicated a failure to represent forest $N_2O$ dynamics, which constitute the largest land use contribution to SSA emissions. FNN was therefore retained as an additional upper-bound uncertainty estimate rather than a primary upscaling model, with the caveat that FNN-based forest estimates should be interpreted with caution. The resulting structural uncertainty range (FNN/XGB/RF) reflects both genuine model uncertainty and, in the case of forests, potential model uncertainty. Comprehensive SSA-wide flux and model comparison statistics are shown in Table S4 and Table S5 in Supporting Information S1.

### 3.2. Important Prediction Variables for $N_2O$ Emissions

The permutation feature importance analysis quantifies each variable's contribution to the models, providing insights into the environmental drivers of $N_2O$ emissions in different ecosystems (Figure 4). To note, these scores

**Table 1**
*Final Model Performance Comparison on the Test Set*

| Land Use | Model | $R^2\uparrow$ | MSE↓ | MAE↓ | ME→ 0 | RMSE↓ | DISO↓ | Mean Flux | Median Flux | SD Flux |
|---|---|---|---|---|---|---|---|---|---|---|
| Grassland | RF | 0.45 | 31.97 | 3.49 | 0.82 | 5.65 | 13.91 | 2.97 | 1.09 | 7.60 |
| | XGB | 0.13 | 50.36 | 4.49 | 0.32 | 7.10 | 14.62 | | | |
| | FNN | 0.55 | 25.83 | 3.03 | −1.44 | 5.08 | 1.71 | | | |
| Forest | RF | 0.36 | 328.65 | 10.55 | −0.25 | 18.13 | 99.15 | 15.30 | 5.55 | 22.58 |
| | XGB | 0.40 | 307.22 | 10.17 | 0.94 | 17.53 | 93.23 | | | |
| | FNN | −0.20 | 610.89 | 11.79 | −1.91 | 24.72 | 1.62 | | | |
| Cropland | RF | 0.36 | 349.61 | 9.08 | 1.53 | 18.70 | 235.89 | 13.70 | 5.42 | 23.45 |
| | XGB | 0.41 | 326.46 | 9.07 | 1.16 | 18.07 | 223.83 | | | |
| | FNN | 0.31 | 377.98 | 8.70 | −4.05 | 19.44 | 1.42 | | | |

*Note.* Flux values in μg $N_2O$-N $m^{-2}$ $h^{-1}$.

indicate the magnitude of variable importance to a model's performance but do not indicate the relationship direction between the variable and $N_2O$ emissions. Negative permutation importance scores can indicate randomly shuffling of the feature's score, suggesting the feature may introduce noise or was overfit in the original model. Here, only RF and XGB driver selections were shown as the FNN used all available variables for training.

Forest ecosystems (Figure 4a) displayed distinct patterns of driver importance in each model. In the XGB model, soil_moisture_3 was the most influential driver (0.685 ± 0.074), followed by soil_temp_3 (0.536 ± 0.061). Within the RF model, soil_temp_3 (0.410 ± 0.051) and soil_moisture_3 (0.219 ± 0.034) were the most influential drivers. The identification of soil_moisture_3 as an important variable in both models (RF: 0.219 ± 0.034; XGB: 0.685 ± 0.074), suggests its consistent influence on forest $N_2O$ emissions. Moreover, month_day_cos, a numerical representation of seasonality, ranked as an important driver in the RF and XGB models (0.146 ± 0.045 and 0.315 ± 0.093), suggesting a seasonal effect on $N_2O$ emissions in SSA forests that are not fully captured by the environmental variables used to train the models.

In grasslands (Figure 4b), the XGB model identified tmin as the dominant driver with an importance score of 0.691 ± 0.119. In the RF model, soil_temp_1 and temp were the most important drivers, with importance scores of 0.273 ± 0.043 and 0.181 ± 0.040, respectively. Rain demonstrated relatively low importance in both models (RF: 0.034 ± 0.034; XGB: 0.028 ± 0.132). In RF, ssrd (−0.023 ± 0.037) and ccov (0.026 ± 0.018) also demonstrated some minor importance, suggesting that vegetation growth played a role in driving $N_2O$ emissions in grasslands.

In croplands (Figure 4c), both models highlighted the importance of fertilization and soil variables on $N_2O$ emissions. Within the XGB model, fert_lvl showed the highest importance (0.451 ± 0.041), followed by soil_temp_3 (0.422 ± 0.059). The RF model identified fert_lvl (0.125 ± 0.027) and soil_temp_3 (0.119 ± 0.016) among the top drivers. The XGB model ranked soil_moisture_3 as a significant predictor (0.272 ± 0.032), while the RF model identified soil_moisture_1 (0.112 ± 0.019) among its important variables. Rain was consistently identified as a relevant predictor in both models (RF: 0.093 ± 0.008; XGB: 0.229 ± 0.032), though its relative importance ranking differed between models.

### 3.3. Regional $N_2O$ Emissions Predictions

In this study, $N_2O$ emissions were predicted and upscaled to the entire SSA region using RF, XGB and FNN models trained on individual SSA sites, using grid cell variable maps from multiple data sets: ERA5 Reanalysis for meteorological and soil data, HaNi for fertilizer information, and ESA World Cover for land use classification. RF estimates are presented as primary results with XGB providing lower-bound sensitivity estimates and FNN providing an additional uncertainty bound (presented as FNN estimate/XGB estimate/RF estimate). Area-based emissions (kg $N_2O$-N $ha^{-1}$ $yr^{-1}$) for all regions and scenarios are provided in Table S6 in Supporting Information S1. SSA-wide and Eastern African $N_2O$ FNN predictions are provided in Figure S9 in Supporting Information S1.

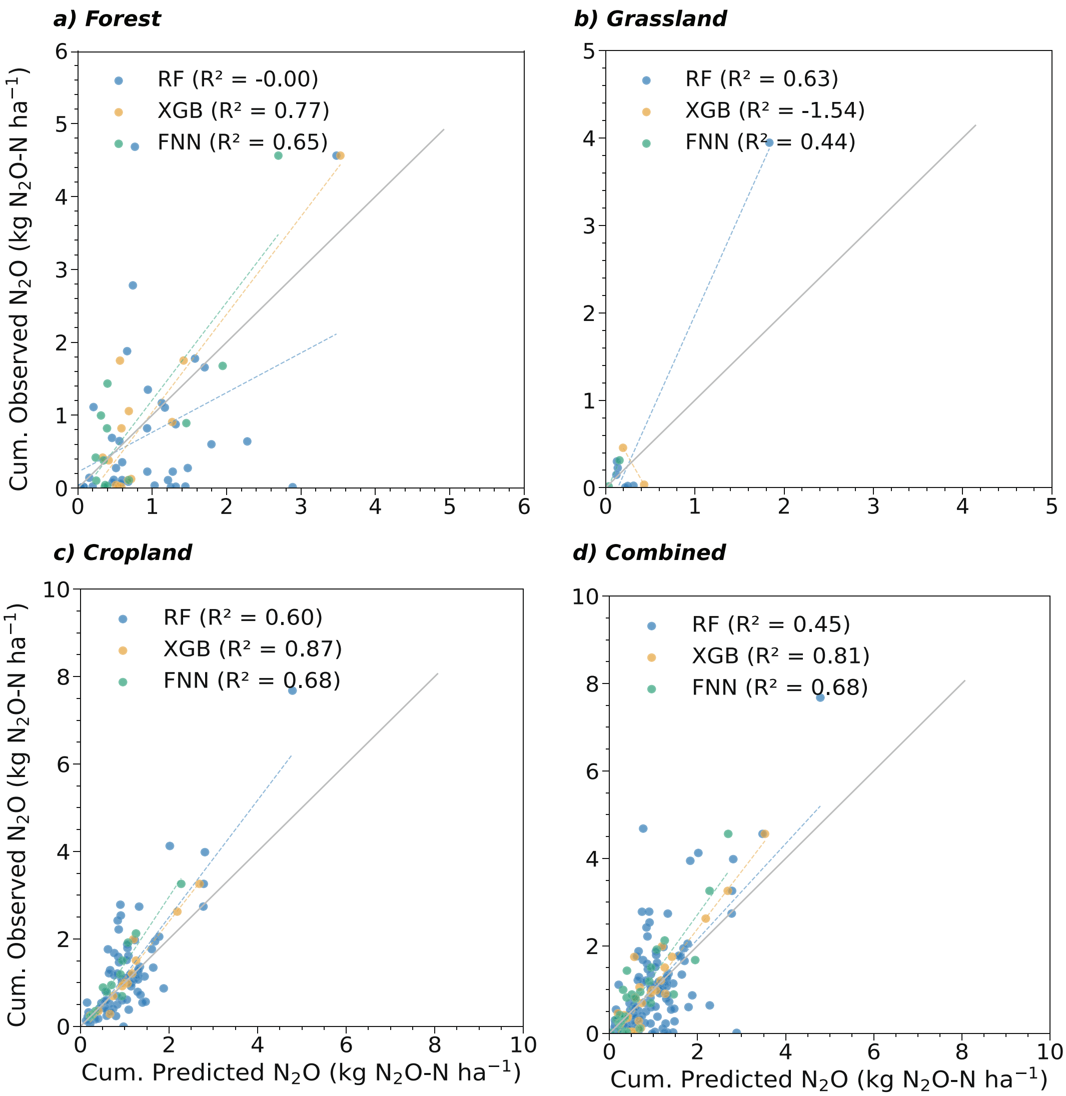


**Figure 3.** Comparison of observed and predicted annual cumulative site $N_2O$ fluxes on the test set: (a) Forest, (b) grassland, (c) cropland, and (d) combined land uses (forest, grassland, and cropland) for each machine learning (ML) model. The individual points represent individual sites and years. For each land use, the ML models were trained on all data sets. Annual cumulative site emissions were calculated by multiplying the mean of all flux measurements by total annual hours (365 × 24) and then converting from μg N $m^{-2}$ to kg N $ha^{-1}$.

Spatial and temporal $N_2O$ predictions showed variable emissions across SSA forests (Figures 5a and 5d). Total $N_2O$ emissions across SSA forests were 119/210/342 Gg N $yr^{-1}$, with Central Africa emerging as the largest regional source (71/138/208 Gg N $yr^{-1}$), followed by Eastern Africa (28/39/83 Gg N $yr^{-1}$), Western Africa (18/31/47 Gg N $yr^{-1}$), and Southern Africa (2/2/4 Gg N $yr^{-1}$).

$N_2O$ emissions from SSA grasslands exhibited considerable regional variation (Figures 5b and 5e), with total emissions of 70/91/132 Gg N $yr^{-1}$. Southern Africa emerged as the dominant source (56/47/65 Gg N $yr^{-1}$), followed by Central Africa (4/18/28 Gg N $yr^{-1}$), Eastern Africa (8/21/28 Gg N $yr^{-1}$), and Western Africa (2/5/12 Gg N $yr^{-1}$).

Total $N_2O$ emissions from croplands were 64/63/64 Gg N $yr^{-1}$ (Figures 5c and 5f), smaller than other land use types. Western Africa was the dominant source (38/22/29 Gg N $yr^{-1}$), followed by Eastern Africa (17/27/26 Gg N $yr^{-1}$), Southern Africa (4/10/4 Gg N $yr^{-1}$), and Central Africa (5/4/4 Gg N $yr^{-1}$).

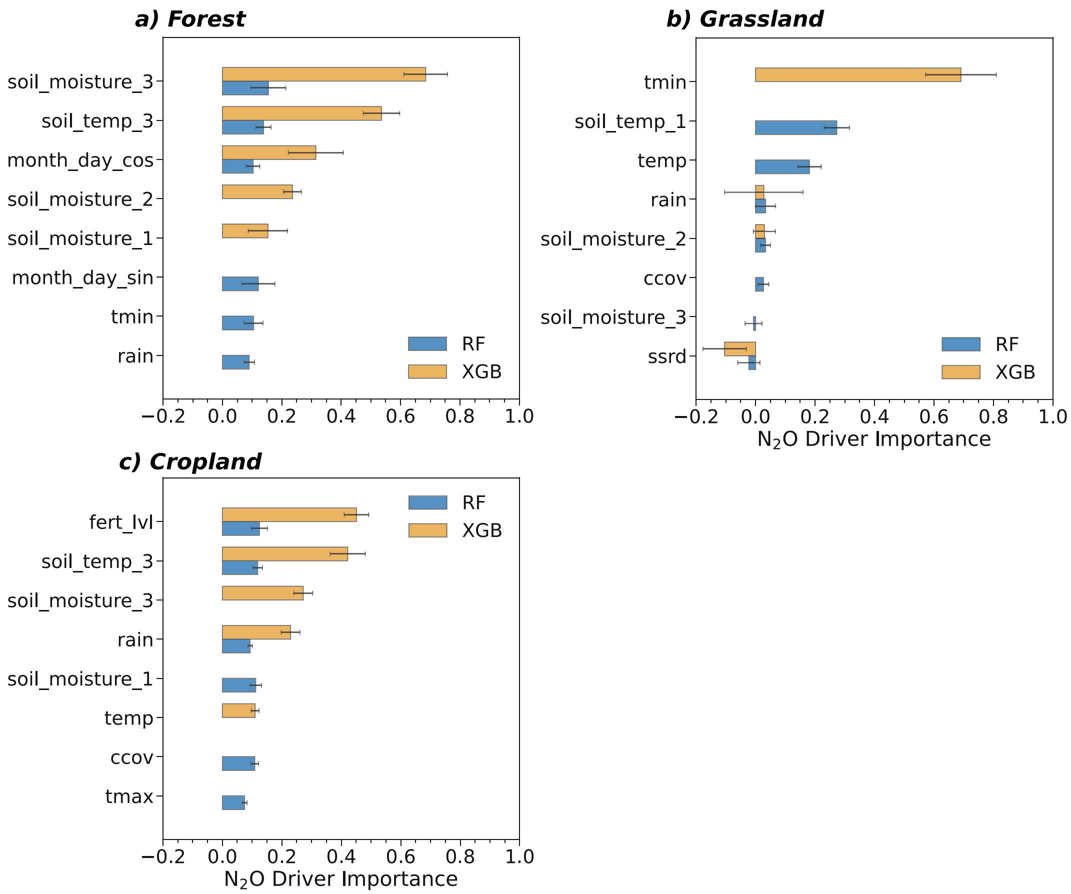


**Figure 4.** Comparison of important drivers of $N_2O$ production from random forest and XGB feature selection processes. (a) Forest, (b) grassland, and (c) cropland. Error bars represent the standard deviation of the permutation importance. Variable defined: soil_temp_1: soil temperature (0–7 cm), soil_temp_3: soil temperature (28–100 cm), temp: air temperature, rain: daily precipitation, soil_moisture_1: soil moisture content (0–7 cm), soil_moisture_2: soil moisture content (7–28 cm), soil_moisture_3: soil moisture content (28–100 cm), month_day_sin: sine transformation of day/month of year (seasonal component), month_day_cos: cosine transformation of day/month of year (seasonal component), tmin: minimum daily temperature, $t_{max}$: maximum daily temperature, ssrd: surface solar radiation downwards, ccov: cloud cover, fert_decay: fertilizer decay rate. For further characterization of these ERA5-Land Reanalysis variables, refer to Table S2 in Supporting Information S1.

### 3.4. Regional $N_2O$ Emissions Predictions Under Climate and Fertilizer Scenarios

In this study, $N_2O$ emissions were predicted and upscaled to the entire SSA region using RF, XGB, and FNN models trained on individual SSA sites for the SSP5-8.5 climate scenario (rain and temperature) and triple N fertilization in SSA croplands, using grid cell variable maps from multiple data sets. RF estimates are presented as primary results with XGB providing lower-bound sensitivity estimates and FNN providing an additional uncertainty bound (presented as FNN estimate/XGB estimate/RF estimate). Area-based emissions (kg $N_2O$-N $ha^{-1}$ $yr^{-1}$) for all regions and scenarios are provided in Table S6 in Supporting Information S1. SSA-wide and Eastern African $N_2O$ FNN predictions are provided in Figure S10 in Supporting Information S1.

SSA forest $N_2O$ predictions under the SSP5-8.5 climate scenario showed notable regional variations compared to baseline (Figures 6a and 6b). Total emissions decreased to 91/106/254 Gg N $yr^{-1}$, representing a 50%–26% (FNN-RF) reduction from baseline. Central African forests maintained their dominance (61/61/139 Gg N $yr^{-1}$), followed by Eastern Africa (17/33/77 Gg N $yr^{-1}$), Western Africa (11/10/33 Gg N $yr^{-1}$), and Southern Africa (1/2/5 Gg N $yr^{-1}$).

$N_2O$ emissions from SSA grasslands under SSP5-8.5 decreased to 19/66/106 Gg N $yr^{-1}$, a 28%–9% (FNN-RF) reduction from baseline (Figures 6c and 6d). Southern Africa maintained dominance (10/34/44 Gg N $yr^{-1}$),

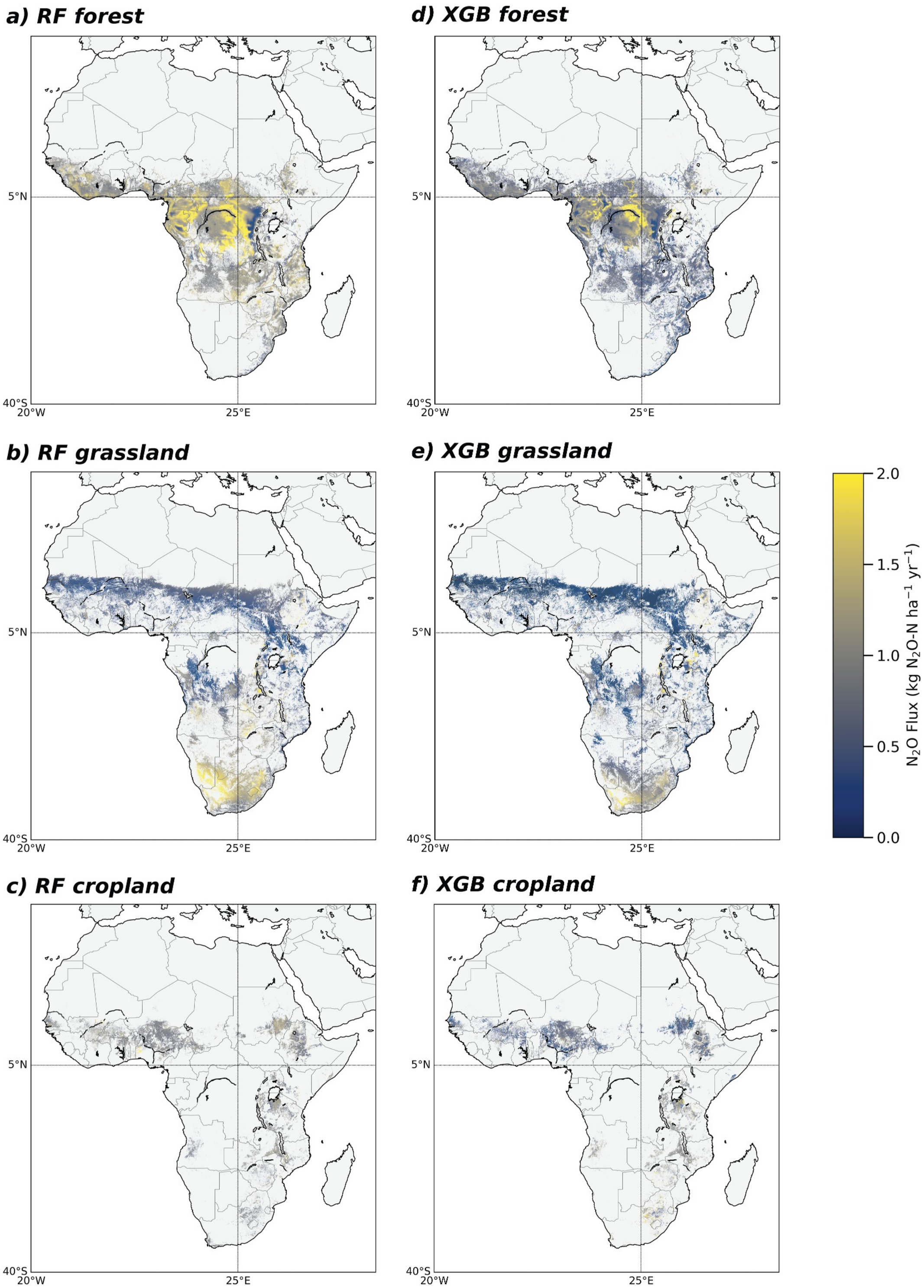


**Figure 5.** Sub-Saharan Africa $N_2O$-N predictions for the baseline using the random forest (RF) and XGBoost (XGB) models for each land use: (a) RF forest, (b) RF grassland, (c) RF cropland, (d) XGB forest, (e) XGB grassland, (f) XGB cropland. For an enhanced view of Eastern Africa, see Figure S7 in Supporting Information S1.

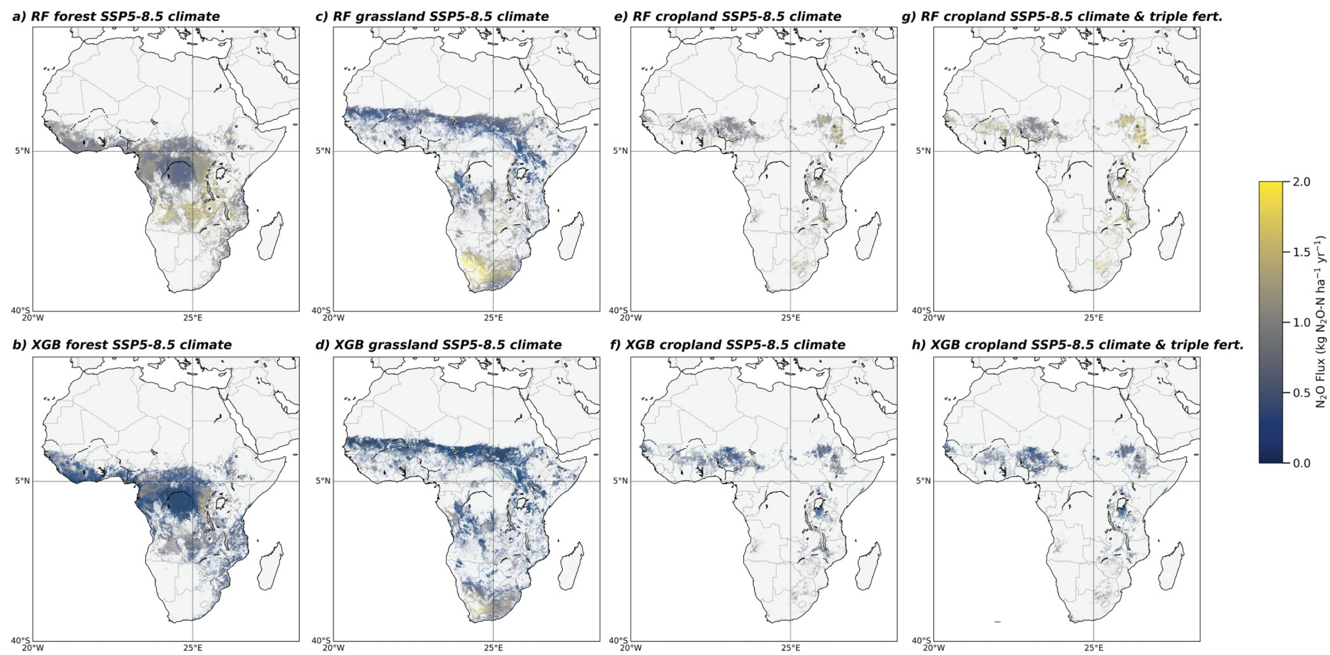


**Figure 6.** Sub-Saharan Africa $N_2O$-N predictions for the SSP5-8.5 climate and triple fertilizer scenario using the random forest (RF) and XGBoost (XGB) models for each land use: (a) RF forest, (b) XGB forest, (c) RF grassland, (d) XGB grassland, (e) RF cropland SSP5-8.5 climate only, (f) XGB cropland SSP5-8.5 climate only, (g) RF cropland SSP5-8.5 climate and triple fertilizer, and (h) XGB cropland SSP5-8.5 climate and triple fertilizer. For an enhanced view of Eastern Africa, see Figure S8 in Supporting Information S1.

followed by Central Africa (2/17/24 Gg N $yr^{-1}$), Eastern Africa (4/12/22 Gg N $yr^{-1}$), and Western Africa (2/3/15 Gg N $yr^{-1}$).

$N_2O$ predictions from SSA croplands showed divergent model responses (Figures 6e and 6f). Under SSP5-8.5 climate-only, emissions ranged from 134/58/76 Gg N $yr^{-1}$, with Western Africa (91/20/31 Gg N $yr^{-1}$) and Eastern Africa (28/27/34 Gg N $yr^{-1}$) as the dominant sources, followed by Southern Africa (4/7/7 Gg N $yr^{-1}$) and Central Africa (12/4/5 Gg N $yr^{-1}$). Under the combined climate-fertilization scenario (Figures 6g and 6h), total emissions increased to 153/67/79 Gg N $yr^{-1}$ from the baseline, with emissions from Eastern Africa (36/32/38 Gg N $yr^{-1}$), Western Africa (99/22/32 Gg N $yr^{-1}$), Southern Africa (6/9/3 Gg N $yr^{-1}$), and Central Africa (12/4/6 Gg N $yr^{-1}$) contributing.

## 4. Discussion

### 4.1. Drivers of $N_2O$ Production

Variable selection identified soil temperature, air temperature and soil moisture as the strongest drivers of $N_2O$ emissions across all land uses, with rainfall showing consistent but lower importance. Other meteorological variables tested, including vpd, patm, ppfd, ccov, and ssrd, showed much lower importance and were not consistently selected across land uses. These findings support previous research across different soil types and ecosystems, as temperature controls microbial activity and decomposition rates that regulate substrate availability for $N_2O$-producing reactions. Soil moisture determines oxygen availability in the soil, controlling $N_2O$ production through nitrification or denitrification pathways (Butterbach-Bahl et al., 2013; Dick et al., 2006; Thornton & Herrero, 2015; Wanyama et al., 2018; Zhuang et al., 2012). The consistent moderate importance of rainfall across all land uses highlights its role in $N_2O$ production, possibly contributing to emission pulses or 'hot moments' (Lawrence & Hall, 2020; Liu et al., 2014). Other known $N_2O$ drivers such as C, N, and pH can influence emissions, but as these variables were not available at sufficient temporal resolution they were not included in the model training.

The importance of soil profile depth varied distinctly between land uses (depending on the model), with surface conditions (0–7 cm) driving grassland emissions while deeper soil layers (28–100 cm) controlled forest and cropland $N_2O$ production. In forests, deeper soil temperature and moisture (28–100 cm) were the dominant environmental drivers, alongside a notable seasonal signal indicated by the importance of month_day_cos. This depth-dependent pattern possibly reflects differences in OC distribution and soil structure between land uses (Pelster et al., 2012; Sessitsch et al., 2001; Silver et al., 2005). While $N_2O$ production is often associated with topsoil processes, the definition of "topsoil" varies across studies, ranging from the surface to 30 cm depth, which overlaps with the deeper soil temperature measurements (Bouwman et al., 1993; Liu et al., 2017; Vilain et al., 2014). Notably, considerable $N_2O$ production can occur at greater depths, with substantial emissions being observed in European spruce forests at 70 cm depths (Goldberg et al., 2008), and at depth in tropical forest soils (Gallarotti et al., 2021; Koehler et al., 2012) suggesting the importance of considering deeper soil layers in $N_2O$ studies. It is important to acknowledge that this study primarily used rainforest environments for training, introducing an ecological bias. The findings may not generalise across all forest ecosystems.

### 4.2. Baseline Regional $N_2O$ Emissions Across SSA

Total annual estimates of $N_2O$ emissions from SSA were 364–538 Gg N $yr^{-1}$ (XGB–RF), with FNN providing an additional lower bound of 253 Gg N $yr^{-1}$, giving a full model uncertainty range of 253–538 Gg N $yr^{-1}$. Forests were the largest contributor (FNN: 47%, XGB: 58%, RF: 64%; 119/210/342 Gg N $yr^{-1}$), followed by grasslands (FNN: 28%, XGB: 16%, RF: 25%; 70/91/132 Gg N $yr^{-1}$), and croplands (FNN: 25%, XGB: 11%, RF: 12% 64/63/64 Gg N $yr^{-1}$). The RF forest emissions (342 Gg N $yr^{-1}$) compare well with Werner et al. (2007), reporting 340 ± 80 Gg N $yr^{-1}$ for tropical rainforests in Africa using the Forest DNDC-Tropica model, while both XGB (210 Gg N $yr^{-1}$) and FNN (119 Gg N $yr^{-1}$) estimates fall outside their uncertainty range. Mean forest emissions (0.39/0.68/1.09 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$, FNN/XGB/RF) aligned well with previous global $N_2O$ modeling studies, which reported ranges of 0.22–0.81 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$ (Zhuang et al., 2012) and 0.65–1.36 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$ (Potter et al., 1996). Similarly, mean grassland emissions (0.40/0.33/0.54 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$, FNN/XGB/RF) were consistent with previous grassland/savanna estimates of 0.21–0.54 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$ (Zhuang et al., 2012) and 0.14–0.65 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$ (Potter et al., 1996). Total cropland emissions (64/63/64 Gg N $yr^{-1}$, FNN/XGB/RF) aligned with SSA estimates (56 Gg N $yr^{-1}$) from Smerald et al. (2023). Cropland emissions averaged 0.84/0.79/0.87 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$ (FNN/XGB/RF), which agreed with Albanito et al. (2017) estimate of 0.8 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$ for Africa and within the range of measured emissions (0.1–4.29 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$) reported in field studies of various SSA countries (Atakora et al., 2019; Kimaro et al., 2016; Mapanda et al., 2011; Rosenstock et al., 2016). None of the reported field studies were used in model training. This shows that RF model predictions align well with both previous $N_2O$ predictions and measured $N_2O$ emissions, though the full model range (approximately ±46%–86% for forests, ±40%–49% for grasslands, and ±3%–49% for croplands including FNN) highlights the substantial structural uncertainty in upscaling from site-level observations, particularly for forests where FNN performed poorly ($R^2 = -0.20$). The high convergence of the three models on total SSA cropland emissions contrasts with the large inter-model spread for forests and grasslands. This likely reflects the more mechanistic role of fertilizer-driven $N_2O$ production in croplands. In natural systems, $N_2O$ production is controlled by complex interactions between environmental conditions such as soil temperature and moisture, which different model architectures may interpret differently. In croplands, however, emissions are more directly tied to N input through the shared fertilizer decay variable, which likely constrains the divergence between model architectures.

In a global context, the prediction totals (253–538 Gg N $yr^{-1}$, FNN-RF) correspond to approximately 1%–3% of the recent global $N_2O$ budget of 17,000 Gg N $yr^{-1}$ (Tian et al., 2024a). When partitioned by emission source, global natural soils contribute 6,400 (3,900–8,600) Gg N $yr^{-1}$, representing more than half of terrestrial $N_2O$ emissions, while direct agricultural emissions contribute approximately 3,900 Gg N $yr^{-1}$ (56% of total anthropogenic emissions) during the 2010s (Tian et al., 2024). The combined SSA estimate represents approximately 4%–8% of global natural soil emissions, suggesting SSA is a modest contributor to the global $N_2O$ budget despite its extensive forest and grassland areas.

At the regional scale, comparison with the most recent global $N_2O$ budget reveals substantial differences in emission magnitudes (Tian et al., 2024). The global budget reports total $N_2O$ emissions for Southern Africa of 610–620 Gg N $yr^{-1}$ and equatorial Africa of 1,380–1,500 Gg N $yr^{-1}$ during the 2010s, partitioned into natural emissions (380 and 980 Gg N $yr^{-1}$, respectively), direct agricultural (70 and 180 Gg N $yr^{-1}$), indirect agricultural

(20 and 50 Gg N $yr^{-1}$), and other anthropogenic sources (190 and 260 Gg N $yr^{-1}$) (Tian et al., 2024). Our regional estimates (FNN/XGB/RF) were: 62/54/73 Gg N $yr^{-1}$ for Southern Africa, 80/160/240 Gg N $yr^{-1}$ for Central Africa, 53/87/137 Gg N $yr^{-1}$ for Eastern Africa, and 58/58/88 Gg N $yr^{-1}$ for Western Africa, are substantially lower than these totals. However, when compared to continent-wide $N_2O$ estimates for Africa from PRIMPA-hist (1,400–1,600 Gg N $yr^{-1}$ between 2000 and 2018) (Mostefaoui et al., 2024), the SSA emissions in this study account for approximately 23%–41% of total African $N_2O$ emissions, excluding industrial sectors (energy, industry, waste) and North Africa, which is much more comparable than to Tian et al. (2024) estimates.

For cropland emissions specifically, our SSA estimate of 63–64 Gg N $yr^{-1}$ represents only 1.6% of global direct agricultural emissions (3,900 Gg N $yr^{-1}$ in 2020), and is considerably lower than those reported by Leitner et al. (2020), who calculated total emissions of 260 Gg N $yr^{-1}$ for maize croplands, and substantially lower than individual African regions in the global budget (70 Gg N $yr^{-1}$ for Southern Africa and 180 Gg N $yr^{-1}$ for equatorial Africa). Leitner et al. (2020) acknowledged that their baseline comparison used EDGAR v5.0 agricultural, primarily an emission database, which represented all crops and managed soils receiving N input, for example, grasslands (10,962,063 $km^2$). Whereas the ESA 2021 World Cover database used in this study is primarily a landcover database, only covering croplands (1,747,673 $km^2$). The discrepancy between cropland $N_2O$ estimates reflects both the different agricultural coverage (crops vs all crops and managed soils) between studies and potentially lower fertilizer application rates in SSA compared to global averages, as SSA agriculture remains predominantly low-input. These comparisons show the uncertainty in regional emission inventories in SSA, where land cover databases show less than 65% spatial agreement (Kerner et al., 2024; Nabil et al., 2020), and highlight the need for cropland emission estimates based on consistent spatial data.

Nevertheless, the spatial patterns of $N_2O$ emissions in this study agree with Leitner et al. (2020), with both investigations identifying the croplands of Eastern Africa as an $N_2O$ emission "hotspot." While this spatial consistency will reflect the concentration of croplands and possibly higher fertilization rates in this region, the agreement between both methodological approaches (ML modeling vs empirical yield gap relationships) provides support for the influence of underlying biogeochemical patterns, validating the regional emission patterns even as questions remain about emission quantification in data-limited regions.

Overall, the spatial patterns of $N_2O$ emissions from natural and agricultural systems in SSA were consistent with previous global modeling studies (Harris et al., 2022; Potter et al., 1996; Stehfest & Bouwman, 2006; Zhuang et al., 2012), identifying Central Africa as a major $N_2O$ source (Figures 5a and 5d). The region's high emissions can be characterized by the environmental conditions present: high temperatures, moisture, and rainfall, combined with clay-rich soils and high OC availability, which promote rapid N cycling and $N_2O$ production (Butterbach-Bahl et al., 2013; Silver et al., 2005; Zhuang et al., 2012). This study also revealed high $N_2O$ emissions in Southern African grasslands, a pattern not observed in previous models (Figures 5b and 5e); however, the absence of observations from this region for model training introduces considerable estimate uncertainties that are not fully quantified in this study. The models were trained predominantly on East African conditions, characterized by bimodal rainfall, more acidic and weathered soils relative to Southern Africa. These differences in climate and soil properties influence soil moisture, microbial activity, and substrate availability, primary drivers of $N_2O$ production, possibly causing the models to misrepresent $N_2O$ dynamics in Southern Africa, with drier conditions, a unimodal rainfall pattern, and distinct soil characteristics (Cardinael et al., 2024). For croplands, our spatial emission patterns align with previous global models (Harris et al., 2022; Potter et al., 1996; Stehfest & Bouwman, 2006; Zhuang et al., 2012), particularly in Western and Eastern Africa.

### 4.3. Climate and Fertilizer Scenarios: Regional $N_2O$ Emissions Across SSA

While climate change is generally expected to increase $N_2O$ emissions through higher air and soil temperatures and changes in rainfall patterns, particularly in drier regions (Butterbach-Bahl & Dannenmann, 2011; Butterbach-Bahl et al., 2013), this study showed different projections. Under the SSP5-8.5 climate-only scenario, the RF model predicted total SSA $N_2O$ emissions would decrease by 19% to 436 Gg N $yr^{-1}$ from the baseline (538 Gg N $yr^{-1}$), with forests remaining the largest contributor (58%; 254 Gg N $yr^{-1}$), followed by grasslands (24%; 106 Gg N $yr^{-1}$), and croplands (17%; 76 Gg N $yr^{-1}$). The XGB model predicted a larger decrease of 37% to 230 Gg N $yr^{-1}$ from baseline (364 Gg N $yr^{-1}$), with forests (46%; 106 Gg N $yr^{-1}$), grasslands (29%; 66 Gg N $yr^{-1}$), and croplands (25%; 58 Gg N $yr^{-1}$). In contrast, the FNN model predicted only a marginal decrease of 4% to

244 Gg N $yr^{-1}$ from baseline (253 Gg N $yr^{-1}$), with markedly different land use contributions: croplands became the dominant contributor (55%; 134 Gg N $yr^{-1}$), while forests (37%; 91 Gg N $yr^{-1}$) and grasslands (8%; 19 Gg N $yr^{-1}$) were substantially lower than at baseline. The near-stable FNN total masks large offsetting changes: a 73% decrease in grasslands (70 → 19 Gg N $yr^{-1}$) and a 23% decrease in forests (119 → 91 Gg N $yr^{-1}$) are almost entirely offset by a 109% increase in croplands (64 → 134 Gg N $yr^{-1}$), driven primarily by Western Africa (91 Gg N $yr^{-1}$). This large FNN cropland projection under future climate should be interpreted with caution, as it likely reflects FNN extrapolating beyond its training distribution under future climate conditions rather than a robust emission signal. The full model range under SSP5-8.5 was therefore 230–436 Gg N $yr^{-1}$ (XGB–RF), with FNN predicting 244 Gg N $yr^{-1}$. The large divergence of cropland predictions under future climate scenarios contrasts with the highly constrained baseline estimates (63–64 Gg N $yr^{-1}$). At baseline conditions, cropland emissions are dominated by N input, a driver shared across the three models. Under future climate scenarios, as temperature, soil moisture, and rainfall extrapolate beyond the training distribution, the different internal representations of environmental driver relationships across model architectures likely become more apparent, resulting in divergent estimates. This highlights a limitation of ML-based upscaling: convergence under observed conditions does not ensure agreement under extrapolated conditions.

Literature on future global $N_2O$ emission predictions is limited, especially for forests and grasslands. Zhuang et al. (2012) found that “large” increases in precipitation, and subsequent soil moisture increases, decreased $N_2O$ emissions. This aligns with experimental evidence showing that higher soil moisture can reduce $N_2O$ production by shifting denitrification toward $N_2$ as the end product (Butterbach-Bahl et al., 2013; Conrad, 1996; Dobbie & Smith, 2003). However, a study of 51 European soils found high $N_2O$ emissions under wetter conditions (>80% water-filled pore space (WFPS), despite the generally accepted optimal soil moisture range of 70%–80% WFPS for $N_2O$ production (Butterbach-Bahl et al., 2013; Davidson et al., 2000; Zechmeister-Boltenstern et al., 2007). The applicability of these findings to SSA remains uncertain due to limited data on optimal soil moisture for $N_2O$ production in (sub-)tropical soils (Almaraz et al., 2020). Temperature effects on $N_2O$ emissions are also complex. Zhuang et al. (2012) found that “small” temperature increases had a greater impact on global $N_2O$ emissions than “medium” increases, possibly due to reduced soil moisture from increased transpiration. Additionally, high temperatures can inhibit denitrification and $N_2O$ production, as demonstrated by Abdalla et al. (2009), who found decreased denitrification activity and $N_2O$ production above 37°C in Irish pasture soils. This study used the SSP5-8.5 scenario, considered the “worst-case” climate projection, with estimated temperature increases of 4.1°C and 4.6°C and rainfall increases of 35% and 10% in boreal winter and summer, respectively, by 2070–2090 (Almazroui et al., 2020).

$N_2O$ emissions from croplands showed divergent model responses to future scenarios, particularly for climate effects. The RF model predicted increases under both scenarios compared to baseline, by 19% in the SSP5-8.5 climate-only scenario (76 Gg N $yr^{-1}$) and by 23% when including both climate and fertilizer changes (79 Gg N $yr^{-1}$). The XGB model predicted an 8% decrease under climate-only (58 Gg N $yr^{-1}$) but a 6% increase when fertilization was included (67 Gg N $yr^{-1}$). The FNN model predicted a much larger increase of 109% under climate-only (134 Gg N $yr^{-1}$) and 139% under the combined scenario (153 Gg N $yr^{-1}$), though as noted above these FNN cropland projections under future climate should be treated with caution. Critically, when isolating the fertilizer effect (comparing climate-only to climate + fertilizer scenarios), RF predicted +4% (76 → 79 Gg N $yr^{-1}$), XGB predicted +16% (58 → 67 Gg N $yr^{-1}$), and FNN predicted +14% (134 → 153 Gg N $yr^{-1}$). While XGB and FNN show reasonable agreement (+14%–16%) in the isolated fertilizer response, the RF estimate (+4%) is more modest, though all three models consistently project that agricultural intensification will increase cropland $N_2O$ emissions.

The RF projections align with Harris et al. (2022), who projected increased global $N_2O$ emissions from anthropogenic sources, including agricultural soils under warming conditions. The modest RF prediction increase in the climate-only scenario suggests that warming and rainfall increases enhanced denitrification activity and $N_2O$ production from existing soil N at current fertilizer rates. The larger increase in the combined climate and fertilizer scenarios across ML models compared to the baseline indicated that N availability currently limits $N_2O$ production in SSA agricultural soils, consistent with Leitner et al. (2020), who also predicted that closing yield gaps through increased fertilizer use in maize croplands would raise $N_2O$ emissions. Estimates from Leitner et al. (2020) were considerably higher than in this study, 1110 ± 190 Gg N $yr^{-1}$ and 1760 ± 230 Gg N $yr^{-1}$ for

50% and 75% yield gap closures, respectively, compared to 67–79 Gg N $yr^{-1}$ (XGB–RF) in this study. As mentioned above, this difference reflects the larger assumed cropland and agricultural area in SSA by the EDGAR database used in their study. When comparing regional emissions per area, the results are more comparable. The RF projection in the SSP5-8.5 climate and fertilizer scenario (1.02 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$) aligns reasonably with the 50% yield gap scenario (1.05 ± 0.19 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$) and is considerably lower than the 75% yield gap value (1.64 ± 0.21 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$). The XGB estimate (0.96 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$) also falls within the range of the 50% yield gap scenario. Both RF and XGB estimates agreed with SSA projections (83 Gg N $yr^{-1}$) from Smerald et al. (2023), using the LDNDC to model a high-emission scenario of increased global fertilizer use. The FNN model projected considerably higher cropland emissions under both future scenarios (134 Gg N $yr^{-1}$ climate-only; 153 Gg N $yr^{-1}$ climate and fertilizer), with area-normalized rates of 1.87 and 2.07 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$, respectively, exceeding the 75% yield gap value of Leitner et al. (2020) (1.64 ± 0.21 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$). These high FNN projections, driven predominantly by Western Africa, may suggest FNN extrapolates poorly beyond its training distribution under future climate conditions, and should therefore be treated as an upper bound that likely overestimates true cropland emissions under the SSP5-8.5 scenario.

Based on the RF model projections, a warmer, wetter climate combined with tripled fertilization rates could increase cropland $N_2O$ emissions across SSA by varying regional percentages. Eastern Africa and Western Africa represent significant emission "hotspots," while Southern Africa also showed substantial increases. In contrast, Central Africa exhibited the lowest emissions increase, likely reflecting a threshold effect, tripling its extremely low fertilizer rates (Figure S13 in Supporting Information S1), may still fall below the threshold needed to increase $N_2O$ emissions. Moreover, for the RF model, the fertilizer response ratio revealed substantial regional variations in how $N_2O$ emissions responds to N inputs under different climate regimes (Table S7 in Supporting Information S1). Eastern Africa showed the strongest synergistic interaction (ratio = 1.69), indicating the $N_2O$ fertilizer response was 69% greater under the SSP5-8.5 climate scenario compared to baseline conditions. Similarly, Southern Africa exhibited a synergistic response (ratio = 1.39). In contrast, Western and Middle Africa displayed slightly antagonistic interactions (ratios of 0.92 and 0.93, respectively), suggesting that climate change may dampen the fertilizer-induced $N_2O$ response in these regions. This highlights the non-linear relationship between fertilizer inputs and $N_2O$ emissions across different regional soil conditions (Hickman et al., 2015; Shcherbak et al., 2014). Overall, the RF estimates of total SSA emission agreed well with Huddell et al. (2020) who estimated that tripling N fertilizer use from 50 to 150 kg N $ha^{-1}$ would increase cropland $N_2O$ emissions in the tropics by 28%. These regional variations highlight the importance of targeted mitigation strategies that account for regional meteorological and soil conditions. However, the climate impacts from increased $N_2O$ emissions would likely be less severe than those from deforestation and subsequent cropland expansion (van Loon et al., 2019).

### 4.4. Climate and Fertilizer Scenarios: Modeling Limitations

Several limitations in the modeling approach should be acknowledged. The future climate scenario SSP5-8.5 was implemented using simple assumptions for a hotter and wetter climate. For example, the relationship between temperature and evaporation/transpiration processes was not considered, which would likely influence $N_2O$ emissions. Moreover, precipitation changes were represented as uniform percentage increases for boreal summer or winter, failing to capture the complexity of precipitation patterns over space and time, though such projections remain uncertain in current climate models. These uniform percentage increases were based on a previous study by Almazroui et al. (2020), estimating the effect of SSP-based scenarios on the climate of SSA. Substantial uncertainties also exist in global soil moisture data sets, which likely affected the $N_2O$ prediction accuracy. The predictions relied on interpolated data from the ERA5 Reanalysis dataset. While this interpolation was necessary due to insufficient variable data in the original data sets, it potentially reduced prediction accuracy since $N_2O$ emissions are highly sensitive to microenvironmental variations across space and time. The model did not consider soil texture, C, N, and pH due to limited temporal resolution or data availability. This exclusion represents a potential source of uncertainty when upscaling $N_2O$ emissions across the entire SSA region. The model did not explicitly include N deposition or manure inputs on pastures and rangeland, which may underestimate emissions across SSA, particularly in future projections. However, these inputs are captured in present-day estimates, as the training data comes from field sites that were subject to atmospheric N deposition and, for grasslands, grazing animal inputs. This assumes these inputs remain constant over time, though the uncertainty in future N deposition and manure trajectories for SSA is likely very large. Morever, explicitly modeling these variables would have also reduced the temporal coverage of the training data: the HaNi dataset spans 1860–2019, while many training observations were collected post-2019, particularly for croplands ($N = 2{,}149$ vs 3,574) and forests ($N = 653$ vs 1584). Lastly, the fertilizer response for croplands was modeled from experimental studies that are generally well-managed, which can lead to higher N use efficiency and lower $N_2O$ emissions, resulting in some potential bias.

The model also likely encountered out-of-sample conditions with higher air and soil temperatures, rainfall, and soil moisture, potentially affecting performance as it predicted based on previously unobserved variables (an edge case handling challenge). To quantify this effect, the SSP5-8.5 conditions were applied to the original training data sets (before the test-train split) using the same methodology as in the upscaling. Then the original trained models for each land use, without retraining them, were used to predict $N_2O$ on these modified data sets, and their performances were evaluated. In the RF models, performance remained relatively stable in forests, croplands, and grasslands, with only slight decreases, no change, or slight increase in $R^2$ values ($R^2 = 0.32$, 0.36, and 0.46, respectively). In the XGB models, forest and cropland models decreased in $R^2$ values ($R^2 = 0.34$ and 0.30), but grassland performance decreased to a much greater extent ($R^2 = -0.19$). This contrasting out-of-sample performance for grasslands was likely due to the limited $N_2O$ measurements available for grasslands ($N = 225$) compared to croplands ($N = 3574$) and forests ($N = 1584$), which likely contributed to greater variability in model performance. To investigate this, a subsampling analysis was performed, where forest and cropland data were randomly subsampled to $N = 225$ and the RF model was retrained 10 times at various sample sizes (225, 500, 1,000, 1,250, 1,500, and full data set). The analysis showed that standard deviations at $N = 225$ were 3–8 times higher than with full data sets (forest: 0.166 vs 0.090; cropland: 0.467 vs 0.055), demonstrating highly unstable performances. As sample size increased, land use types also showed improvements in mean performance ($r > 0.84$). While grassland estimates should be interpreted with caution, this finding highlights the need for more high-resolution $N_2O$ measurements in SSA grasslands. While the extent of out-of-sample conditions may have varied among models and land use type, the subsampling analysis demonstrates that data scarcity was likely a driver of poor model performance for grassland predictions under future climate scenarios.

Further investigation into the decreased $N_2O$ emissions in both forests and grasslands, using individual model runs for temperature and rainfall scenarios, revealed that rising temperatures were the primary driver, while rainfall effects were minimal (Figure S11 in Supporting Information S1). The FNN grassland model was an exception, showing a large increase in $N_2O$ emissions in the rainfall-only scenario and a large decrease in the temperature-only scenario; however, when combined in the SSP5-8.5 scenario, overall $N_2O$ emissions still decreased (Figure S11 in Supporting Information S1). This highlights the dominance of temperature variables as emission predictors in the ML models. Moreover, this study aimed to include seasonal $N_2O$ flux estimates; however, this was not possible because the model did not adequately account for different controlling factors during the rainy and dry seasons. The models sometimes predicted higher $N_2O$ fluxes during the dry season compared to the rainy season, contradictory to typical expectations. This decrease in forest and grassland $N_2O$ emissions and the inadequate seasonal fluxes likely resulted from the observed temperature sensitivity of the ML models and an overrepresentation of training sites where rainfall or soil moisture is not the primary constraint on $N_2O$ production.

### 4.5. SSA Data Characterization: Model Performance

The variable model performance across land uses likely reflects structural differences in the underlying data sets, particularly in sampling frequency, data set size, and flux variability, rather than the land uses themselves. Grassland was the most data-limited land use, comprising only 4 sites and 225 observations, with a mean inter-measurement gap of 11.7 days and approximately 29% of sampling intervals exceeding one week. Despite this, grassland fluxes were characterized by low mean emissions at three of four sites (mean: 1.2–1.6 μg $N_2O$-N $m^{-2}$ $h^{-1}$), resulting in a relatively homogeneous emissions regime that may have been more manageable for machine-learning models. The exception, site 18IW-CHESP-GL (mean: 21.7 μg $N_2O$-N $m^{-2}$ $h^{-1}$), suggests that between-site variability (defined as the standard deviation of site mean fluxes: 10.15 μg $N_2O$-N $m^{-2}$ $h^{-1}$) rather than within-site dynamics was the dominant source of variance in grasslands. Forest data sets were characterized by sparse and irregular sampling (mean gap: 11.2 days; 34% of intervals exceeding 1 week), as well as high between-site flux variability. Site means ranged from 0.23 to 39.3 μg $N_2O$-N $m^{-2}$ $h^{-1}$, with between-site variability

(10.44 μg $N_2O$-N $m^{-2}$ $h^{-1}$) approaching within-site variability (defined as the mean of per-site standard deviations; 11.01 μg $N_2O$-N $m^{-2}$ $h^{-1}$). This near-parity suggests that forest flux dynamics were driven by site-specific factors, likely reflecting differences in soil type, tree species, management across the 8 sites, and the $N_2O$ sampling regime, which limits the ability of models to generalize across sites. In particular, the high-flux outlier site 24RD-Yang (mean: 39.3, SD: 27.7 μg $N_2O$-N $m^{-2}$ $h^{-1}$) likely affected the FNN forest model ($R^2 = -0.20$). FNN are sensitive to outliers when training data are sparse (Vishwakarma et al., 2021), whereas tree-based methods such as RF and XGB are more robust to outliers, consistent with their relatively stronger performances for forests. Cropland had the largest data set (14 sites, 3,574 observations) and the highest within-site flux variability (within-site SD: 17.19 μg $N_2O$-N $m^{-2}$ $h^{-1}$), likely driven by fertilization- and rainfall-induced emission pulses. Although cropland sampling was on average more frequent (mean gap: 6.7 days), two sites (23AS-DTC and 23AS-UZF) had mean inter-measurement gaps exceeding 27 days, indicating that episodic fluxes were likely undersampled in the training data. The consistently moderate $R^2$ values across all three models for cropland (0.31–0.41) suggest that no model was able to fully resolve the high flux, short-duration emission events characteristic of managed agricultural soils under these sampling regimes. These results suggest that model selection should be guided mainly by sampling density and between-site homogeneity of the available data set. FNN may be preferable where sampling is frequent and sites have homogeneous fluxes, while RF and XGB offers greater robustness under sparse and heterogeneous conditions. Detailed data set characteristics, including measured variable coverage and sampling protocols, are provided in Section S1.0 in Supporting Information S1.

### 4.6. Outlook and Recommendations for $N_2O$ Measurements in SSA

$N_2O$ fluxes exhibit high variability across small temporal and spatial scales owing to microbial production pathways, controlled by multiple meteorological and soil parameters, including soil and air temperature, soil moisture content, rainfall, pH, and redox potential. Thus, $N_2O$ emissions are inherently challenging to characterize, particularly in SSA, where there is limited research infrastructure and restricted access to high-resolution instrumentation. Despite these challenges, expanding $N_2O$ flux measurements across temporal and spatial scales is crucial for better understanding soil-atmospheric exchanges across climatic, vegetational, and agricultural environments in SSA. This is particularly important for high-resolution measurements as "hot moment" emissions likely dominate $N_2O$ fluxes. Moreover, standardizing measurement designs is essential to improve statistical validation when scaling from plot to regional levels.

Currently, $N_2O$ flux measurements in SSA primarily rely on manual closed chambers, offering ease of use, cost-effectiveness, and experimental flexibility. However, temporal coverage is low due to the reliance on manual measurements, typically limited to weekly or monthly measurements. This frequency often misses $N_2O$ "hot moments" during specific events, such as fertilization or rainfall events, leading to considerable uncertainty in SSA emission estimates. Automated chamber systems connected to precise $N_2O$ analyzers (e.g., cavity ring-down spectroscopy or tunable diode laser instruments) can address temporal limitations. To better understand $N_2O$ emission patterns, high-resolution meteorological and soil parameter data are essential, particularly for improving predictions from ML models. Currently, these measurements are seldom taken or are taken relatively sporadically; therefore, this study had to rely on interpolated data from global models, reducing model prediction accuracy since $N_2O$ production is sensitive to microclimatic changes and can show delayed responses to environmental events (Rautakoski et al., 2024). Moreover, the coverage of SSA environments remains poor, especially in grasslands and forests (mainly rainforest sites in Western and Central Africa). Currently, grassland and cropland measurements are concentrated in specific regions, particularly Eastern Africa (Kenya and Tanzania—Figure 1). This geographic bias affects $N_2O$ predictions in ML models, as Eastern African meteorological and soil conditions do not represent SSA. Soil physiology differs significantly across the continent based on SOC, soil texture (silt and clay content), and mineral soil composition (von Fromm et al., 2021). The wetter regions of East Africa, where most studies take place, have more acidic and weathered soils compared to drier regions in Southern Africa. Seasonality also varies considerably, with East Africa typically experiencing bimodal rainfall with two distinct rainy seasons (enabling two growing seasons and higher fertilizer inputs), while West and Southern Africa generally experience unimodal rainfall with a single rainy season per year, and Central Africa has a distinct bimodal rain pattern, generally receiving high precipitation throughout the year (Longandjo & Rouault, 2024; Palmer et al., 2023). Since the ML models were trained predominantly on data from East African, it may systematically misrepresent $N_2O$ dynamics in other SSA regions. These soil and climatic differences strongly influence soil moisture, oxygen availability, microbial activity, and substrate availability, primary

10.1029/2026GB009141


drivers of $N_2O$ production, likely reducing the accuracy of upscaled $N_2O$ estimates for regions with substantially different conditions.

Short-term recommendations for improved $N_2O$ measurements would be: (a) Implementing daily measurements, preferably over a full year, particularly in croplands, where fallow periods have been shown to release considerable $N_2O$ under certain conditions (Verhoeven et al., 2017); (b) Meteorological and soil measurements should be conducted alongside $N_2O$ sampling, with air temperature, soil temperature, soil moisture, and rainfall prioritised as standard parameters. When possible, soil (mineral) N and C measurements should also be conducted at a reduced frequency. In resource-constrained settings without automated systems, manual measurements should occur three to four times weekly, with increased frequency following specific events such as fertilisation or rainfall to capture emission pulses (Garland et al., 2014; Kennedy et al., 2013; Verhoeven & Six, 2014). Diurnal cycles could be examined to find the most representative time for measurements. Furthermore, air temperature, soil temperature, and soil moisture measurements, using accessible sensors, should be taken alongside these $N_2O$ measurements (e.g., Decock et al. (2015)).

Long-term strategies should focus on developing comprehensive monitoring networks. This includes establishing an expanded network of long-term monitoring stations across diverse SSA land use types, emphasising high-resolution monitoring of $N_2O$ emissions and environmental parameters with automated systems, including chamber-based measurements. Existing networks include INDAAF (https://indaaf.obs-mip.fr/) and FLUXNET Eddy Covariance infrastructure (www.fluxnet.org); however, these networks currently do not measure $N_2O$. INDAAF measures certain trace gases passively (e.g., $SO_2$, $HNO_3$, $NH_3$, $NO_2$, $O_3$), while the existing FLUXNET Eddy Covariance infrastructure is limited to $CO_2$ and water vapour measurements. Both provide an opportunity to incorporate and expand $N_2O$ monitoring. Moreover, their databases serve as an example for developing standardised protocols for chamber-based measurement, where a community-driven platform would facilitate accessible databases for $N_2O$ measurements and associated environmental parameters. Moreover, this platform could provide standardised measurement protocols that outline procedures for data collection that can be applied across research sites, tailored to the challenges of measurements in less developed regions. Additionally, user-friendly data management protocols should be developed to maintain data quality standards and accessible storage. Long-term $N_2O$ monitoring programmes require investment in local expertise and capacity building, including training programmes, technical support networks, and collaborative research initiatives. International partnerships can foster two-way knowledge transfer while ensuring research is relevant to local needs, particularly in agricultural studies. Addressing these research challenges in SSA is crucial to improve measurements, enhancing predictive modelling capabilities and our understanding of $N_2O$ processes in (sub-)tropical environments, which remain understudied in current literature. Success in these efforts will contribute considerably to global greenhouse gas monitoring and climate change mitigation strategies.

However, a successful SSA monitoring network must extend beyond improving our understanding of $N_2O$ processes in (sub-)tropical environments to encompass the broader socio-economic and institutional elements for effective implementation. Stakeholder engagement strategies should prioritise knowledge exchange with farmers, policymakers, and local communities. Farmer engagement provides a "reality-check" for mitigation strategies through real-world application while creating educational platforms where scientific insights and traditional agricultural knowledge can be integrated. This ensures that monitoring protocols are practically feasible and appropriate for diverse SSA farming practices and systems. Policymaker engagement is equally crucial for establishing supportive regulatory frameworks and designing evidence-based mitigation policies, while providing feedback mechanisms to assess policy effectiveness in reducing emissions. Additionally, successful network implementation requires sustainable financing models that combine international funding with local investment, institutional capacity building to ensure long-term data continuity. These transdisciplinary approaches, spanning technical, social, and economic dimensions, are essential for transforming monitoring networks from research tools into drivers for sustainable agricultural transformation and climate change mitigation across SSA.

## 5. Conclusion

SSA needs to boost agricultural production for regional and global food security in coming decades. Increasing the currently low fertilizer use would enhance productivity but could increase regional $N_2O$ emissions, though large uncertainties exist in emission estimates, particularly under climate change. Three ML models (RF, XGB, FNN) were trained using previous $N_2O$ measurements to predict emissions across SSA for cropland, forests, and grasslands. RF was selected for primary upscaling results, with XGB providing a lower-bound and FNN an additional uncertainty estimate. $N_2O$ emissions were most sensitive to soil and air temperature, soil moisture, and rainfall, with fertilizer decay important for croplands. Annual estimates totaled 253–538 Gg N $yr^{-1}$ (FNN–RF models; 1%–3% of natural and anthropogenic global emissions (Tian et al., 2024)), with forests as the largest contributor (119–342 Gg N $yr^{-1}$, 0.39–1.09 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$, FNN–RF), followed by grasslands (70–132 Gg N $yr^{-1}$, 0.33–0.54 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$, FNN–RF) and croplands (63–64 Gg N $yr^{-1}$, 0.79–0.87 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$, XGB–RF)). The full model range (±3%–86% depending on land use and model) reflects structural uncertainty in upscaling from site-level observations to regional scales. Under the SSP5-8.5 climate scenario, the RF model predicted total SSA $N_2O$ emissions would decrease by 19% (to 436 Gg N $yr^{-1}$) due to increased temperatures and rainfall effects, with forests at 254 Gg N $yr^{-1}$ (0.90 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$), and grasslands at 106 Gg N $yr^{-1}$ (0.39 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$). The XGB model predicted larger decreases (37%–230 Gg N $yr^{-1}$). The FNN model predicted a 4% decrease in $N_2O$ (to 244 Gg N $yr^{-1}$). Reductions in soil $N_2O$ emissions from forests and grasslands were mostly offset by a significant increase in $N_2O$ emissions from croplands. However, these FNN cropland projections under a future climate should be interpreted with caution. In the combined climate and fertilizer scenario, cropland emissions increased to 67–79 Gg N $yr^{-1}$ (XGB–RF; 0.96–1.15 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$) or 153 Gg N $yr^{-1}$ (FNN; 2.07 kg $N_2O$-N $ha^{-1}$ $yr^{-1}$) as an upper bound. Despite greater warming and rainfall predicted for SSA, natural ecosystems showed decreased emissions, possibly due to high temperatures increasing evaporation/transpiration resulting in drier soils. Meanwhile, croplands show large emission increases under the climate and tripling fertilizer scenario, suggesting SSA will become a larger agricultural emitter while overall continental emissions may decrease. Improving model prediction accuracy requires higher temporal resolution measurements of $N_2O$ and environmental factors (especially temperature, soil moisture, and rainfall) from more sites, with better geographic representation. Moreover, precipitation variability remains uncertain in climate models, which generates further uncertainty in $N_2O$ predictions. Long-term climate observation strategies in SSA should focus on improving research infrastructure for $N_2O$ measurements with standardised protocols and an accessible regional data repository. Beyond model accuracy, upscaling to regional emission estimates also depends on reliable land use area estimates, particularly croplands.

## Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

## Availability Statement

The consolidated forest, grassland, and cropland data sets used in this study are available in Barthel et al. (2022); Daelman et al. (2024); Gütlein et al. (2018); Hickman et al. (2014), Hickman et al. (2015); Laub et al. (2023); Shumba et al. (2023); Wanyama et al. (2018); Zheng et al. (2019); Wachiye et al. (2020). The consolidated data set bundle is publicly available through the PANGAEA® Data Publisher (Agredazywczuk et al., 2026). The data set from Lourenço et al. (2026) will be available from 1 June 2027 (Zenodo, CC BY-NC 4.0 license) due to restrictions by Nestlé (third-party funder) on the timing of data publication. All ML model training and SSA upscaling Python scripts are publicly available through Zenodo (Agredazywczuk, 2026).).

**Acknowledgments**

This research was supported by the Swiss National Science Foundation (SNSF) under Grant 200021_207348. For R. Daelman's work, the CongoFlux team's help maintaining all the equipment should be acknowledged. The main funder of R. Dealman's work was the Special Research Fund of Ghent University, through the Methusalem mandate of Pascal Boeckx (project no. BOF.MET.2021.0004.01). This paper contains modified Copernicus Climate Change Service information 2025. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. For R. Balokah's work, Brigitte Veronica Cuendet and Elijah Adarkwa should be acknowledged. For K. Lourenço's work, part of the data set used in this study, specifically the data from Côte d'Ivoire, originates from the Ground Zero Project, led jointly by Wageningen University & Research and the Nestlé Institute of Agricultural Sciences and funded by Nestlé. Open access publishing facilitated by Universitat Bern, as part of the Wiley - Universitat Bern agreement via the Consortium Of Swiss Academic Libraries.

## References

Abdalla, M., Jones, M., Smith, P., & Williams, M. (2009). Nitrous oxide fluxes and denitrification sensitivity to temperature in irish pasture soils. *Soil Use and Management*, *25*(4), 376–388. https://doi.org/10.1111/j.1475-2743.2009.00237.x

Aderele, M. O., Srivastava, A. K., Butterbach-Bahl, K., & Rahimi, J. (2025). Integrating machine learning with agroecosystem modelling: Current state and future challenges. *European Journal of Agronomy*, *168*, 127610. https://doi.org/10.1016/j.eja.2025.127610

Agredazywczuk, P. (2026). pzywczuk/n2o-ssa-models: Nitrous oxide emissions across Sub-Saharan Africa: Meta-analysis and data-driven modelling ($n_2$o-ssa-models) [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.19740369

Agredazywczuk, P., Ouma, T., Barthel, M., Otinga, A., Njoroge, R., Butterbach-Bahl, K., et al. (2026). Nitrous oxide emissions across Sub-Saharan Africa: Meta-analysis and data-driven modelling [dataset bundled publication] [Dataset]. *PANGAEA*. https://doi.org/10.1594/PANGAEA.987298

Albanito, F., Lebender, U., Cornulier, T., Sapkota, T. B., Brentrup, F., Stirling, C. M., & Hillier, J. (2017). Direct nitrous oxide emissions from tropical and sub-tropical agricultural systems–A review and modelling of emission factors. *Scientific Reports*, *7*(1), 44235. https://doi.org/10.1038/srep44235

Almaraz, M., Wong, M. Y., & Yang, W. H. (2020). Looking back to look ahead: A vision for soil denitrification research. *Ecology*, *101*(1), e02917. https://doi.org/10.1002/ecy.2917

Almazroui, M., Saeed, F., Saeed, S., Islam, M. N., Ismail, M., Klutse, N. A. B., & Siddiqui, M. H. (2020). Projected change in temperature and precipitation over Africa from cmip6. *Earth Systems and Environment*, *4*(3), 455–475. https://doi.org/10.1007/s41748-020-00161-x

Atakora, W. K., Kwakye, P. K., Weymann, D., & Brüggemann, N. (2019). Stimulus of nitrogen fertilizers and soil characteristics on maize yield and nitrous oxide emission from ferric luvisol in the Guinea savanna agro-ecological zone of Ghana. *Scientific African*, *6*, e00141. https://doi.org/10.1016/j.sciaf.2019.e00141

Baggs, E. M., Smales, C. L., & Bateman, E. J. (2010). Changing ph shifts the microbial source as well as the magnitude of $n_2o$ emission from soil. *Biology and Fertility of Soils*, *46*(8), 793–805. https://doi.org/10.1007/s00374-010-0484-6

Barthel, M., Bauters, M., Baumgartner, S., Bodé, S., Kenber, J., Geeraert, L., et al. (2022). Low $N_2O$ and variable $CH_4$ fluxes from tropical forest soils of the Congo Basin. *Nature Communications*, *13*(1), 330. https://doi.org/10.1038/s41467-022-27978-6

Barton, L., Wolf, B., Rowlings, D., Scheer, C., Kiese, R., Grace, P., et al. (2015). Sampling frequency affects estimates of annual nitrous oxide fluxes. *Scientific Reports*, *5*(1), 15912. https://doi.org/10.1038/srep15912

Basche, A., Miguez, F., Kaspar, T., & Castellano, M. (2014). Do cover crops increase or decrease nitrous oxide emissions? A meta-analysis. *Journal of Soil and Water Conservation*, *69*(6), 471–482. https://doi.org/10.2489/jswc.69.6.471

Berardi, D., Brzostek, E., Blanc-Betes, E., Davison, B., DeLucia, E. H., Hartman, M. D., et al. (2020). 21st-century biogeochemical modeling: Challenges for century-based models and where do we go from here? *GCB Bioenergy*, *12*(10), 774–788. https://doi.org/10.1111/gcbb.12730

Bouwman, A., Fung, I., Matthews, E., & John, J. (1993). Global analysis of the potential for $n_2o$ production in natural soils. *Global Biogeochemical Cycles*, *7*(3), 557–597. https://doi.org/10.1029/93gb01186

Bouwman, A. F., Boumans, L. J. M., & Batjes, N. H. (2002). Emissions of $n_2o$ and no from fertilized fields: Summary of available measurement data. *Global Biogeochemical Cycles*, *16*(4), 6-1–13. https://doi.org/10.1029/2001GB001811

Braker, G., & Conrad, R. (2011). Advances in Applied Microbiology, *Chapter 2 - Diversity, structure, and size of $n_2o$-producing microbial communities in Soils—What matters for their functioning?* In A. I. Laskin, S. Sariaslani, & G. M. Gadd (Eds.), Vol. 75, 33–70. Academic Press. https://doi.org/10.1016/B978-0-12-387046-9.00002-5

Butterbach-Bahl, K., Baggs, E. M., Dannenmann, M., Kiese, R., & Zechmeister-Boltenstern, S. (2013). Nitrous oxide emissions from soils: How well do we understand the processes and their controls? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1621), 20130122. https://doi.org/10.1098/rstb.2013.0122

Butterbach-Bahl, K., & Dannenmann, M. (2011). Denitrification and associated soil $n_2o$ emissions due to agricultural activities in a changing climate. *Current Opinion in Environmental Sustainability*, *3*(5), 389–395. https://doi.org/10.1016/j.cosust.2011.08.004

Cardinael, R., Barton, L., Corbeels, M., Six, J., Rowlings, D., Shumba, A., et al. (2024). Soil n2o emissions during dry fallow periods. *Global Change Biology*, *30*(7), e17403. https://doi.org/10.1111/gcb.17403

Charles, A., Rochette, P., Whalen, J. K., Angers, D. A., Chantigny, M. H., & Bertrand, N. (2017). Global nitrous oxide emission factors from agricultural soils after addition of organic amendments: A meta-analysis. *Agriculture, Ecosystems and Environment*, *236*, 88–98. https://doi.org/10.1016/j.agee.2016.11.021

Chemura, A., Gleixner, S., & Gornott, C. (2024). Dataset of the suitability of major food crops in Africa under climate change. *Scientific Data*, *11*(1), 294. https://doi.org/10.1038/s41597-024-03118-1

Chen, T., & Guestrin, C. (2016). Xgboost. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785

Conrad, R. (1996). Soil microorganisms as controllers of atmospheric trace gases ($H_2$, co, $ch_4$, ocs, $n_2o$, and no). *Microbiological Reviews*, *60*(4), 609–640. https://doi.org/10.1128/mr.60.4.609-640.1996

Costa, C., Galford, G. L., Coe, M. T., Macedo, M., Jankowski, K., O'Connell, C., & Neill, C. (2021). Modeling nitrous oxide emissions from large-scale intensive cropping systems in the southern Amazon. *Frontiers in Sustainable Food Systems*, *5*, 701416. https://doi.org/10.3389/fsufs.2021.701416

Daelman, R., Bauters, M., Barthel, M., Bulonza, E., Lefevre, L., Mbifo, J., et al. (2024). Spatiotemporal variability of $CO_2$, $N_2O$ and $CH_4$ fluxes from a semi-deciduous tropical forest soil in the Congo Basin. *EGUsphere. (Preprint)*. https://doi.org/10.5194/egusphere-2024-2346

Davidson, E. A., Keller, M., Erickson, H. E., Verchot, L. V., & Veldkamp, E. (2000). Testing a conceptual model of soil emissions of nitrous and nitric oxides. *BioScience*, *50*(8), 667–680. https://doi.org/10.1641/0006-3568(2000)050[0667:TACMOS]2.0.CO;2

Decock, C., Lee, J., Necpalova, M., Pereira, E. I. P., Tendall, D. M., & Six, J. (2015). Mitigating $n_2o$ emissions from soil: From patching leaks to transformative action. *SOIL*, *1*(2), 687–694. https://doi.org/10.5194/soil-1-687-2015

Dick, J., Skiba, U., Munro, R., & Deans, D. (2006). Effect of n-fixing and non n-fixing trees and crops on no and $n_2o$ emissions from Senegalese soils. *Journal of Biogeography*, *33*(3), 416–423. https://doi.org/10.1111/j.1365-2699.2005.01421.x

Dobbie, K. E., & Smith, K. A. (2003). Nitrous oxide emission factors for agricultural soils in Great Britain: The impact of soil water-filled pore space and other controlling variables. *Global Change Biology*, *9*(2), 204–218. https://doi.org/10.1046/j.1365-2486.2003.00563.x

Ehrhardt, F., Soussana, J.-F., Bellocchi, G., Grace, P., McAuliffe, R., Recous, S., et al. (2018). Assessing uncertainties in crop and pasture ensemble model simulations of productivity and $n_2o$ emissions. *Global Change Biology*, *24*(2), e603–e616. https://doi.org/10.1111/gcb.13965

Falconnier, G. N., Cardinael, R., Corbeels, M., Baudron, F., Chivenge, P., Couedel, A., et al. (2023). The input reduction principle of agroecology is wrong when it comes to mineral fertilizer use in Sub-Saharan Africa. *Outlook on Agriculture*, *52*(3), 311–326. https://doi.org/10.1177/00307270231199795

FAO. (2021). The state of food and agriculture 2021: Making agrifood systems more resilient to shocks and stresses. *Food and Agriculture Organization of the United Nations*. Retrieved from https://openknowledge.fao.org/server/api/core/bitstreams/7e6336f8-d90d-4936-805b-f612a218f0c8/content

Fao, U. W. W., & IFAD. (2018). *The state of food security and nutrition in the world 2018: Building climate resilience for food security and nutrition*. Food and Agriculture Organisation of the United Nations. Retrieved from https://www.fao.org/3/I9553EN/i9553en.pdf

Gaillard, R. K., Jones, C. D., Ingraham, P., Collier, S., Izaurralde, R. C., Jokela, W., et al. (2018). Underestimation of $n_2o$ emissions in a comparison of the daycent, dndc, and epic models. *Ecological Applications*, *28*(3), 694–708. https://doi.org/10.1002/eap.1674

Gallarotti, N., Barthel, M., Verhoeven, E., Pereira, E. I. P., Bauters, M., Baumgartner, S., et al. (2021). In-depth analysis of $n_2o$ fluxes in tropical forest soils of the Congo Basin combining isotope and functional gene analysis. *The ISME Journal*, *15*(11), 3357–3374. https://doi.org/10.1038/s41396-021-01004-x

Garland, G. M., Suddick, E., Burger, M., Horwath, W., & Six, J. (2014). Direct $n_2o$ emissions from a mediterranean vineyard: Event-related baseline measurements. *Agriculture, Ecosystems and Environment*, *195*, 44–52. https://doi.org/10.1016/j.agee.2014.05.018

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., et al. (2010). Food security: The challenge of feeding 9 billion people. *Science*, *327*(5967), 812–818. https://doi.org/10.1126/science.1185383

Goldberg, S. D., Knorr, K.-H., & Gebauer, G. (2008). $N_2o$ concentration and isotope signature along profiles provide deeper insight into the fate of $n_2o$ in soils. *Isotopes in Environmental and Health Studies*, *44*(4), 377–391. https://doi.org/10.1080/10256010802507433

Groffman, P. M., Butterbach-Bahl, K., Fulweiler, R. W., Gold, A. J., Morse, J. L., Stander, E. K., et al. (2009). Challenges to incorporating spatially and temporally explicit phenomena (hotspots and hot moments) in denitrification models. *Biogeochemistry*, *93*(1–2), 49–77. https://doi.org/10.1007/s10533-008-9277-5

Grosso, S. J. D., Parton, W. J., Mosier, A. R., Hartman, M. D., Brenner, J., & Ojima, D. S. (2006). Simulated interaction of carbon dynamics and nitrogen trace gas fluxes using the daycent model1.

Gütlein, A., Gerschlauer, F., Kikoti, I., & Kiese, R. (2018). Impacts of climate and land use on $N_2O$ and $CH_4$ fluxes from tropical ecosystems in the Mt. Kilimanjaro region, Tanzania. *Global Change Biology*, *24*(3), 1239–1255. https://doi.org/10.1111/gcb.13944

Hamrani, A., Akbarzadeh, A., & Madramootoo, C. A. (2020). Machine learning for predicting greenhouse gas emissions from agricultural soils. *Science of The Total Environment*, *741*, 140338. https://doi.org/10.1016/j.scitotenv.2020.140338

Harris, E., Yu, L., Wang, Y. P., Mohn, J., Henne, S., Bai, E., et al. (2022). Warming and redistribution of nitrogen inputs drive an increase in terrestrial nitrous oxide emission factor. *Nature Communications*, *13*(1), 4310. https://doi.org/10.1038/s41467-022-32001-z

Hickman, J. E., Havlikova, M., Kroeze, C., & Palm, C. A. (2011). Current and future nitrous oxide emissions from African agriculture. *Current Opinion in Environmental Sustainability*, *3*(5), 370–378. https://doi.org/10.1016/j.cosust.2011.08.001

Hickman, J. E., Palm, C. A., Mutuo, P., Melillo, J. M., & Tang, J. (2014). Nitrous oxide ($n_2o$) emissions in response to increasing fertilizer addition in maize (zea mays l.) agriculture in Western Kenya. *Nutrient Cycling in Agroecosystems*, *100*(2), 177–187. https://doi.org/10.1007/s10705-014-9636-7

Hickman, J. E., Tully, K. L., Groffman, P. M., Diru, W., & Palm, C. A. (2015). A potential tipping point in tropical agriculture: Avoiding rapid increases in nitrous oxide fluxes from agricultural intensification in Kenya. *Journal of Geophysical Research: Biogeosciences*, *120*(5), 938–951. https://doi.org/10.1002/2015JG002913

Hodnett, M., & Tomasella, J. (2002). Marked differences between Van genuchten soil water-retention parameters for temperate and tropical soils: A new water-retention pedo-transfer functions developed for tropical soils. *Geoderma*, *108*(3), 155–180. https://doi.org/10.1016/S0016-7061(02)00105-2

Huddell, A. M., Galford, G. L., Tully, K. L., Crowley, C., Palm, C. A., Neill, C., et al. (2020). Meta-analysis on the potential for increasing nitrogen losses from intensifying tropical agriculture. *Global Change Biology*, *26*(3), 1668–1680. https://doi.org/10.1111/gcb.14951

IPCC. (2006). *2006 ipcc guidelines for national greenhouse gas inventories. Kanagawa*. IGES.

IPCC. (2023). In H. Lee & J. Romero (Eds.), *Climate change 2023: Synthesis report. Contribution of working groups I, II and III to the sixth assessment report of the intergovernmental panel on climate change*. IPCC. https://doi.org/10.59327/IPCC/AR6-9789291691647

Jayne, T. S., & Sanchez, P. A. (2021). Agricultural productivity must improve in Sub-Saharan Africa. *Science*, *372*(6546), 1045–1047. https://doi.org/10.1126/science.abf5413

Kennedy, T. L., Suddick, E. C., & Six, J. (2013). Reduced nitrous oxide emissions and increased yields in California tomato cropping systems under drip irrigation and fertigation. *Agriculture, Ecosystems and Environment*, *170*, 16–27. https://doi.org/10.1016/j.agee.2013.02.002

Kerner, H., Nakalembe, C., Yang, A., Zvonkov, I., McWeeny, R., Tseng, G., & Becker-Reshef, I. (2024). How accurate are existing land cover maps for agriculture in Sub-Saharan Africa? *Scientific Data*, *11*(1), 486. https://doi.org/10.1038/s41597-024-03306-z

Khalil, M., AlSayed, A., Liu, Y., & Vanrolleghem, P. A. (2023). Machine learning for modeling $n_2o$ emissions from wastewater treatment plants: Aligning model performance, complexity, and interpretability. *Water Research*, *245*, 120667. https://doi.org/10.1016/j.watres.2023.120667

Kim, D., Hernandez-Ramirez, G., & Giltrap, D. (2013). Linear and nonlinear dependency of direct nitrous oxide emissions on fertilizer nitrogen input: A meta-analysis. *Agriculture, Ecosystems and Environment*, *168*, 53–65. https://doi.org/10.1016/j.agee.2012.02.021

Kim, D., Thomas, A. D., Pelster, D., Rosenstock, T. S., & Sanz-Cobena, A. (2016). Greenhouse gas emissions from natural ecosystems and agricultural lands in Sub-Saharan Africa: Synthesis of available data and suggestions for further research. *Biogeosciences*, *13*(16), 4789–4809. https://doi.org/10.5194/bg-13-4789-2016

Kim, T., Jin, Z., Smith, T. M., Liu, L., Yang, Y., Yang, Y., et al. (2021). Quantifying nitrogen loss hotspots and mitigation potential for individual fields in the Us corn belt with a metamodeling approach. *Environmental Research Letters*, *16*(7), 075008. https://doi.org/10.1088/1748-9326/ac0d21

Kimaro, A. A., Mpanda, M., Rioux, J., Aynekulu, E., Shaba, S., Thiong'o, M., et al. (2016). Is conservation agriculture "climate-smart" for maize farmers in the highlands of Tanzania? Nutrient cycling. *Agro-Ecosystems*, *105*(3), 217–228. https://doi.org/10.1007/s10705-015-9711-8

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Koehler, B., Corre, M. D., Steger, K., Well, R., Zehe, E., Sueta, J. P., & Veldkamp, E. (2012). An in-depth look into a tropical lowland forest soil: Nitrogen-addition effects on the contents of $n_2o$, $co_2$ and $ch_4$ and $n_2o$ isotopic signatures Down to 2-m depth. *Biogeochemistry*, *111*(1), 695–713. https://doi.org/10.1007/s10533-012-9711-6

Kravchenko, A. N., & Robertson, G. P. (2015). Statistical challenges in analyses of chamber-based soil $co_2$ and $n_2o$ emissions data. *Soil Science Society of America Journal*, *79*(1), 200–211. https://doi.org/10.2136/sssaj2014.08.0325

Laub, M., Corbeels, M., Mathu Ndungu, S., Mucheru-Muna, M. W., Mugendi, D., Necpalova, M., et al. (2023). Combining manure with mineral n fertilizer maintains maize yields: Evidence from four long-term experiments in Kenya. *Field Crops Research*, *291*, 108788. https://doi.org/10.1016/j.fcr.2022.108788

Laub, M., Necpalova, M., Van de Broek, M., Corbeels, M., Ndungu, S. M., Mucheru-Muna, M. W., et al. (2024). Modeling integrated soil fertility management for maize production in Kenya using a bayesian calibration of the daycent model. *Biogeosciences*, *21*(16), 3691–3716. https://doi.org/10.5194/bg-21-3691-2024

Lawrence, N. C., & Hall, S. J. (2020). Capturing temporal heterogeneity in soil nitrous oxide fluxes with a robust and low-cost automated chamber apparatus. *Atmospheric Measurement Techniques*, *13*(7), 4065–4078. https://doi.org/10.5194/amt-13-4065-2020

Lee, J.-Y., Marotzke, J., Bala, G., Cao, L., Corti, S., Dunne, J., et al. (2021). Future global climate: Scenario-based projections and near-term information. In V. Masson-Delmotte, P. Zhai, & A. Pirani (Eds.), *Climate change 2021: The physical science basis. Contribution of working groups I, II and III to the sixth assessment report of the intergovernmental panel on climate change* (pp. 553–672). https://doi.org/10.1017/9781009157896.006

Leitner, S., Pelster, D. E., Werner, C., Merbold, L., Baggs, E. M., Mapanda, F., & Butterbach-Bahl, K. (2020). Closing maize yield gaps in Sub-Saharan Africa will boost soil $n_2o$ emissions. *Current Opinion in Environmental Sustainability*, *47*, 95–105. https://doi.org/10.1016/j.cosust.2020.08.018

Levy, P., Mobbs, D., Jones, S., Milne, R., Campbell, C., & Sutton, M. (2007). Simulation of fluxes of greenhouse gases from European grasslands using the dndc model. *Agriculture, Ecosystems and Environment*, *121*(1), 186–192. https://doi.org/10.1016/j.agee.2006.12.019

Li, C., Frolking, S., & Frolking, T. A. (1992). A model of nitrous oxide evolution from soil driven by rainfall events: 1. Model structure and sensitivity. *Journal of Geophysical Research*, *97*(D9), 9759–9776. https://doi.org/10.1029/92JD00509

Liu, X., Qi, Y., Dong, Y., Peng, Q., He, Y., Sun, L., et al. (2014). Response of soil $n_2o$ emissions to precipitation pulses under different nitrogen availabilities in a semiarid temperate steppe of Inner Mongolia, China. *Journal of Arid Land*, *6*(4), 410–422. https://doi.org/10.1007/s40333-013-0211-x

Liu, X., Zhang, Q., Li, S., Zhang, L., & Ren, J. (2017). Simulated $nh_4^+$-n deposition inhibits $ch_4$ uptake and promotes $n_2o$ emission in the meadow steppe of Inner Mongolia, China. *Pedosphere*, *27*(2), 306–317. https://doi.org/10.1016/S1002-0160(17)60318-7

Longandjo, G.-N. T., & Rouault, M. (2024). Revisiting the seasonal cycle of rainfall over Central Africa. *Journal of Climate*, *37*(3), 1015–1032. https://doi.org/10.1175/JCLI-D-23-0281.1

Lourenço, K. S., Barthel, M., Miguel-Angel, L. M., Ouattara, K. L., Six, J., Giller, K., et al. (2026). Nitrous oxide ($n_2o$) flux measurement dataset from sub-saharan African croplands [dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.20638539

Lowder, S. K., Skoet, J., & Raney, T. (2016). The number, size, and distribution of farms, smallholder farms, and family farms worldwide. *World Development*, *87*, 16–29. https://doi.org/10.1016/j.worlddev.2015.10.041

Mapanda, F., Wuta, M., Nyamangara, J., & Rees, R. M. (2011). Effects of organic and mineral fertilizer nitrogen on greenhouse gas emissions and plant-captured carbon under maize cropping in Zimbabwe. *Plant and Soil*, *343*(1–2), 67–81. https://doi.org/10.1007/s11104-011-0753-7

Mostefaoui, M., Ciais, P., McGrath, M. J., Peylin, P., Patra, P. K., & Ernst, Y. (2024). Greenhouse gas emissions and their trends over the last 3 decades across Africa. *Earth System Science Data*, *16*(1), 245–275. https://doi.org/10.5194/essd-16-245-2024

Mueller, N. D., Gerber, J. S., Johnston, M., Ray, D. K., Ramankutty, N., & Foley, J. A. (2012). Closing yield gaps through nutrient and water management. *Nature*, *490*(7419), 254–257. https://doi.org/10.1038/nature11420

Nabil, M., Zhang, M., Bofana, J., Wu, B., Stein, A., Dong, T., et al. (2020). Assessing factors impacting the spatial discrepancy of remote sensing based cropland products: A case study in Africa. *International Journal of Applied Earth Observation and Geoinformation*, *85*, 102010. https://doi.org/10.1016/j.jag.2019.102010

Ntinyari, W., Reichel, R., Gweyi-Onyango, J. P., Giweta, M., Wissel, H., Masso, C., et al. (2023). Nitrogen fertilization and liming increased $co_2$ and $n_2o$ emissions from tropical ferralsols, but not from a vertisol. *Soil Use and Management*, *39*(3), 1125–1139. https://doi.org/10.1111/sum.12927

Palmer, P. I., Wainwright, C. M., Dong, B., Maidment, R. I., Wheeler, K. G., Gedney, N., et al. (2023). Drivers and impacts of eastern African rainfall variability. *Nature Reviews Earth and Environment*, *4*, 254–270. https://doi.org/10.1038/s43017-023-00451-1

Pelster, D. E., Chantigny, M. H., Rochette, P., Angers, D. A., Rieux, C., & Vanasse, A. (2012). Nitrous oxide emissions respond differently to mineral and organic nitrogen sources in contrasting soil types. *Atmospheric Pollutant and Trace Gas*, *41*(2), 427–435. https://doi.org/10.2134/jeq2011.0261

Philibert, A., Loyce, C., & Makowski, D. (2013). Prediction of $n_2o$ emission from local information with random forest. *Environmental Pollution*, *177*, 156–163. https://doi.org/10.1016/j.envpol.2013.02.019

Portmann, R. W., Daniel, J. S., & Ravishankara, A. R. (2012). Stratospheric ozone depletion due to nitrous oxide: Influences of other gases. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1593), 1256–1264. https://doi.org/10.1098/rstb.2011.0377

Potter, C. S., Matson, P. A., Vitousek, P. M., & Davidson, E. A. (1996). Process modeling of controls on nitrogen trace gas emissions from soils worldwide. *Journal of Geophysical Research*, *101*(D1), 1361–1377. https://doi.org/10.1029/95JD02028

Rautakoski, H., Korkiakoski, M., Mäkelä, J., Koskinen, M., Minkkinen, K., Aurela, M., et al. (2024). Exploring temporal and spatial variation of nitrous oxide flux using several years of peatland forest automatic chamber data. *Biogeosciences*, *21*(7), 1867–1886. https://doi.org/10.5194/bg-21-1867-2024

Rochette, P., Liang, C., Pelster, D., Bergeron, O., Lemke, R., Kroebel, R., et al. (2018). Soil nitrous oxide emissions from agricultural soils in Canada: Exploring relationships with soil, crop and climatic variables. *Agriculture, Ecosystems and Environment*, *254*, 69–81. https://doi.org/10.1016/j.agee.2017.10.021

Rosenstock, T. S., Mpanda, M., Pelster, D. E., Butterbach-Bahl, K., Rufino, M. C., Thiong'o, M., et al. (2016). Greenhouse gas fluxes from agricultural soils of Kenya and Tanzania. *Journal of Geophysical Research: Biogeosciences*, *121*(6), 1568–1580. https://doi.org/10.1002/2016JG003341

Rurinda, J., Zingore, S., Jibrin, J. M., Balemi, T., Masuki, K., Andersson, J. A., et al. (2020). Science-based decision support for formulating crop fertilizer recommendations in Sub-Saharan Africa. *Agricultural Systems*, *180*, 102790. https://doi.org/10.1016/j.agsy.2020.102790

Saha, D., Basso, B., & Robertson, G. P. (2021). Machine learning improves predictions of agricultural nitrous oxide ($n_2o$) emissions from intensively managed cropping systems. *Environmental Research Letters*, *16*(2), 024004. https://doi.org/10.1088/1748-9326/abd2f3

Sessitsch, A., Weilharter, A., Gerzabek, M. H., Kirchmann, H., & Kandeler, E. (2001). Microbial population structures in soil particle size fractions of a long-term fertilizer field experiment. *Applied and Environmental Microbiology*, *67*(9), 4215–4224. https://doi.org/10.1128/AEM.67.9.4215-4224.2001

Shcherbak, I., Millar, N., & Robertson, G. P. (2014). Global metaanalysis of the nonlinear response of soil nitrous oxide ($n_2o$) emissions to fertilizer nitrogen. *Proceedings of the National Academy of Sciences*, *111*(25), 9199–9204. https://doi.org/10.1073/pnas.1322434111

Sheahan, M., & Barrett, C. B. (2017). Ten striking facts about agricultural input use in Sub-Saharan Africa. *Food Policy*, *67*, 12–25. https://doi.org/10.1016/j.foodpol.2016.09.010

Shumba, A., Chikowo, R., Corbeels, M., Six, J., Thierfelder, C., & Cardinael, R. (2023). Long-term tillage, residue management and crop rotation impacts on $n_2o$ and $ch_4$ emissions from two contrasting soils in sub-humid Zimbabwe. *Agriculture, Ecosystems and Environment*, *341*, 108207. https://doi.org/10.1016/j.agee.2022.108207

Silver, W. L., Thompson, A. W., McGroddy, M., Varner, R. K., Dias, J. D., Silva, H. P. d., & Keller, M. (2005). Fine root dynamics and trace gas fluxes in two lowland tropical forest soils. *Global Change Biology*, *11*(2), 290–306. https://doi.org/10.1111/j.1365-2486.2005.00903.x

Six, J., Feller, C., Denef, K., Ogle, S., de Moraes Sa, J. C., & Albrecht, A. (2002). Soil organic matter, biota and aggregation in temperate and tropical soils - Effects of no-tillage. *Agronomie*, *22*(7–8), 755–775. https://doi.org/10.1051/agro:2002043

Smerald, A., Kraus, D., Rahimi, J., Fuchs, K., Kiese, R., Butterbach-Bahl, K., & Scheer, C. (2023). A redistribution of nitrogen fertiliser across global croplands can help achieve food security within environmental boundaries. *Communications Earth and Environment*, *4*(1), 315. https://doi.org/10.1038/s43247-023-00970-8

Soci, C., Hersbach, H., Simmons, A., Poli, P., Bell, B., Berrisford, P., et al. (2024). The era5 global reanalysis from 1940 to 2022. *Quarterly Journal of the Royal Meteorological Society*, *150*(764), 4014–4048. https://doi.org/10.1002/qj.4803

Stehfest, E., & Bouwman, L. (2006). $N_2O$ and NO emission from agricultural fields and soils under natural vegetation: Summarizing available measurement data and modeling of global annual emissions. *Nutrient Cycling in Agroecosystems*, *74*(3), 207–228. https://doi.org/10.1007/s10705-006-9000-7

Stursa, D., & Dolezel, P. (2019). Comparison of relu and linear saturated activation functions in neural network for universal approximation. In *2019 22nd international conference on process control* (pp. 146–151). https://doi.org/10.1109/PC.2019.8815057

Syakila, A., & Kroeze, C. (2011). The global nitrous oxide budget revisited. *Greenhouse Gas Measurement and Management*, *1*(1), 17–26. https://doi.org/10.3763/ghgmm.2010.0007

Thornton, P. K., & Herrero, M. (2015). Adapting to climate change in the mixed crop and livestock farming systems in Sub-Saharan Africa. *Nature Climate Change*, *5*(9), 830–836. https://doi.org/10.1038/nclimate2754

Tian, H., Bian, Z., Shi, H., Qin, X., Pan, N., Lu, C., et al. (2022). HaNi: A historical dataset of anthropogenic nitrogen inputs to the terrestrial biosphere (pp. 1860–2019). https://doi.org/10.1594/PANGAEA.942069

Tian, H., Pan, N., Thompson, R., Canadell, P., Suntharalingam, P., Regnier, P., et al. (2024). Global nitrous oxide budget (1980–2020). *Earth System Science Data*, *16*(6), 2543–2604. https://doi.org/10.5194/essd-16-2543-2024

Tomasella, J., & Hodnett, M. (2004). Pedotransfer functions for tropical soils. In *Development of pedotransfer functions in soil hydrology* (Vol. 30, pp. 415–429). Elsevier. https://doi.org/10.1016/S0166-2481(04)30021-8

Tully, K. L., Hickman, J. E., Russo, T. A., Neill, C., Matata, P., Nyadzi, G., et al. (2023). The fate of nitrogen during agricultural intensification in East Africa: Nitrogen budgets in contrasting agroecosystems. *Journal of Geophysical Research: Biogeosciences*, *128*(7), e2022JG007128. https://doi.org/10.1029/2022JG007128

United Nations. (2019). *World population prospects 2019: Highlights*. United Nations. Retrieved from https://population.un.org/wpp/Publications/(ST/ESA/SER.A/423

Van Groenigen, J. W., Velthof, G. L., Oenema, O., Van Groenigen, K. J., & Van Kessel, C. (2010). Towards an agronomic assessment of $n_2o$ emissions: A case study for arable crops. *European Journal of Soil Science*, *61*(6), 903–913. https://doi.org/10.1111/j.1365-2389.2009.01217.x

van Ittersum, M. K., Van Bussel, L. G. J., Wolf, J., Grassini, P., Van Wart, J., Guilpart, N., et al. (2016). Can Sub-Saharan Africa feed itself? *Proceedings of the National Academy of Sciences*, *113*(52), 14964–14969. https://doi.org/10.1073/pnas.1610359113

Van Koppen, B. (2003). Water reform in Sub-Saharan Africa: What is the difference? *Physics and Chemistry of the Earth, Parts A/B/C*, *28*(20), 1047–1053. https://doi.org/10.1016/j.pce.2003.08.022

van Loon, M. P., Hijbeek, R., ten Berge, H. F. M., De Sy, V., ten Broeke, G. A., Solomon, D., & Van Ittersum, M. K. (2019). Impacts of intensifying or expanding cereal cropping in Sub-Saharan Africa on greenhouse gas emissions and food security. *Global Change Biology*, *25*(11), 3720–3730. https://doi.org/10.1111/gcb.14783

Velthof, G., Oudendag, D., & Oenema, O. (2007). *Development and application of the integrated nitrogen model miterra-europe* (Vol. 102). Alterra Report. Alterra, Wageningen.

Verhoeven, E., Pereira, E., Decock, C., Garland, G., Kennedy, T., Suddick, E., et al. (2017). $N_2o$ emissions from California farmlands: A review. *California Agriculture*, *71*(3), 148–159. https://doi.org/10.3733/ca.2017a0026

Verhoeven, E., & Six, J. (2014). Biochar does not mitigate field-scale $n_2o$ emissions in a northern California vineyard: An assessment across two years. *Agriculture, Ecosystems and Environment*, *191*, 27–38. https://doi.org/10.1016/j.agee.2014.03.008

Vilain, G., Garnier, J., Decuq, C., & Lugnot, M. (2014). Nitrous oxide production from soil experiments: Denitrification prevails over nitrification. *Nutrient Cycling in Agroecosystems*, *98*(2), 169–186. https://doi.org/10.1007/s10705-014-9604-2

Vishwakarma, G. K., Paul, C., & Elsawah, A. (2021). A hybrid feedforward neural network algorithm for detecting outliers in non-stationary multivariate time series. *Expert Systems with Applications*, *184*, 115545. https://doi.org/10.1016/j.eswa.2021.115545

Vitousek, P. M. (1984). Litterfall, nutrient cycling, and nutrient limitation in tropical forests. *Ecology*, *65*(1), 285–298. https://doi.org/10.2307/1939481

von Fromm, S. F., Hoyt, A. M., Lange, M., Acquah, G. E., Aynekulu, E., Berhe, A. A., et al. (2021). Continental-scale controls on soil organic carbon across Sub-Saharan Africa. *SOIL*, *7*(1), 305–332. https://doi.org/10.5194/soil-7-305-2021

Vuichard, N., Soussana, J., Ciais, P., Viovy, N., Ammann, C., Calanca, P., et al. (2007). Estimating the greenhouse gas fluxes of European grasslands with a process-based model: 1. Model evaluation from in situ measurements. *Global Biogeochemical Cycles*, *21*(1), GB1004. https://doi.org/10.1029/2005GB002611

Wachiye, S., Merbold, L., Vesala, T., Rinne, J., Räsänen, M., Leitner, S., & Pellikka, P. (2020). Soil greenhouse gas emissions under different land-use types in savanna ecosystems of Kenya. *Biogeosciences*, *17*(8), 2149–2167. https://doi.org/10.5194/bg-17-2149-2020

Wan, Y., Chen, F., Fan, L., Sun, D., He, H., Dai, Y., & Chen, Y. (2023). Conversion of surface $ch_4$ concentrations from gosat satellite observations using xgboost algorithm. *Atmospheric Environment*, *301*, 119694. https://doi.org/10.1016/j.atmosenv.2023.119694

Wang, X., Qin, Y., Wang, Y., Xiang, S., & Chen, H. (2019). Reltanh: An activation function with vanishing gradient resistance for sae-based dnns and its application to rotating machinery fault diagnosis. *Neurocomputing*, *363*, 88–98. https://doi.org/10.1016/j.neucom.2019.07.017

Wanyama, I., Pelster, D., Arias-Navarro, C., Butterbach-Bahl, K., Verchot, L., & Rufino, M. (2018). Management intensity controls soil $n_2o$ fluxes in an Afromontane ecosystem. *Science of The Total Environment*, *624*, 769–780. https://doi.org/10.1016/j.scitotenv.2017.12.081

Werner, C., Butterbach-Bahl, K., Haas, E., Hickler, T., & Kiese, R. (2007). A global inventory of $n_2o$ emissions from tropical rainforest soils using a detailed biogeochemical model. *Global Biogeochemical Cycles*, *21*(3). https://doi.org/10.1029/2006GB002909

Wiggins, S. (2009). Can the smallholder model deliver poverty reduction and food security for a rapidly growing population in Africa? *Food and Agriculture Organisation of the United Nations (FAO)*

Wohl, E., Barros, A., Brunsell, N., Chappell, N. A., Coe, M., Giambelluca, T., et al. (2012). The hydrology of the humid tropics. *Nature Climate Change*, *2*(9), 655–662. https://doi.org/10.1038/nclimate1556

Xu, Y., & Cai, Z. (2007). Denitrification characteristics of subtropical soils in China affected by soil parent material and land use. *European Journal of Soil Science*, *58*(6), 1293–1303. https://doi.org/10.1111/j.1365-2389.2007.00923.x

Xu, Y., Xu, Z., Cai, Z., & Reverchon, F. (2013). Review of denitrification in tropical and subtropical soils of terrestrial ecosystems. *Journal of Soils and Sediments*, *13*(4), 699–710. https://doi.org/10.1007/s11368-013-0650-1

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., et al. (2022). ESA WorldCover 10 m 2021 v200. https://doi.org/10.5281/zenodo.7254221

Zechmeister-Boltenstern, S., Schaufler, G., & Kitzler, B. (2007). No, $no_2$, $n_2o$, $co_2$ and $ch_4$ fluxes from soils under different land use: Temperature sensitivity and effects of soil moisture. In *Geophysical research abstracts* (Vol. 9).07968

Zheng, J., Qu, Y., Kilasara, M. M., Mmari, W. N., & Funakawa, S. (2019). Soil-atmosphere exchange of nitrous oxide in two Tanzanian croplands: Effects of nitrogen and stover management. *Agricultural and Forest Meteorology*, *275*, 24–36. https://doi.org/10.1016/j.agrformet.2019.05.009

Zhu, Y., Merbold, L., Leitner, S., Xia, L., Pelster, D. E., Diaz-Pines, E., et al. (2020). Influence of soil properties on $n_2o$ and $co_2$ emissions from excreta deposited on tropical pastures in Kenya. *Soil Biology and Biochemistry*, *140*, 107636. https://doi.org/10.1016/j.soilbio.2019.107636

Zhuang, Q., Lu, Y., & Chen, M. (2012). An inventory of global $N_2O$ emissions from the soils of natural terrestrial ecosystems. *Atmospheric Environment*, *47*, 66–75. https://doi.org/10.1016/j.atmosenv.2011.11.036