Agriculture sector:

Estimation of Direct Nitrous Oxide (N_2O) Emissions from Synthetic Fertilizers

Prateek Sharma[1,2,*] and Bruno Basso[1]
1) Department of Earth and Environmental Sciences, Michigan State University, 2) Climate TRACE
**Email- sharm165@msu.edu*

1. Introduction

Nitrous Oxide (N_2O) is a highly potent greenhouse gas (GHG), with a global warming potential (GWP) of 298 that of CO_2 on a 100-year timescale (EPA, 2018). The concentration of atmospheric N_2O has increased by more than 20%, from 270 parts per billion (ppb) in 1750 to 331 ppb in 2018 (Tian et al., 2020). This dramatic shift is primarily driven by increased anthropogenic sources that raise current total global emissions to ~17 teragram (Tg) nitrogen (N) (Syakila and Kroeze et al., 2011; Thompson et al., 2019; Tian et al., 2020). Approximately 52% of anthropogenic emissions come from the direct emissions from Nitrogen (N) application in the agricultural sector (Tian et al., 2020). Direct N_2O emissions from soils are primarily produced during two microbial-driven biological processes, nitrification and denitrification. Nitrification is the aerobic microbial oxidation, in which ammonium ion is oxidized into nitrate and N_2O is released as a byproduct.

On the other hand, denitrification is the anaerobic microbial (mainly bacterial) reduction of nitrate to nitrite and then to the gasses NO, N_2O, and N_2. N_2O production depends on the amount of mineral N substrates in the soil, i.e., ammonium and nitrate. Therefore, additions of mineral N fertilizers and other sources of N (manures, residue) to agricultural soil are considered the primary drivers of N_2O emissions and higher atmospheric concentrations. Lassaletta et al. (2014) estimated that only 47% of the reactive nitrogen added globally onto cropland is converted into harvested products, and the rest is lost into the environment. For sustainable agriculture, it is essential to quantify these emissions with more confidence and find ways to mitigate climate change by reducing them. Northup et al. (2021) have shown in their recently published work that there is a potential to reduce 71% (1,744 kg CO_2e/ha) of synthetic fertilizer's greenhouse gas emissions through a combination of innovations in digital agriculture, crop and microbial genetics, and electrification (using electrical farm equipment instead of fuel-based) in the next 15 years. According to this study, process-based models guided via high-resolution monitoring systems can be used to optimize the N fertilizer use for the crop, which can further reduce the N_2O emissions. Similarly, customizing crop genetics and root traits can help improve the plant's nitrogen use efficiency. The addition of N-fixing microbes (in addition to those that form root nodules) in the soil can reduce the need for chemical fertilizer, eventually leading to reduced emissions.

Various efforts have been made in the past to quantify N_2O emissions at the regional and global levels from the agricultural sector. The FAOSTAT emissions database of the Food and Agriculture Organization of the United Nations (FAO) covers emissions of N_2O from agriculture by country and globally from 1961 to 2018 (Tubiello et al., 2021). The United Nations Framework Convention on Climate Change (UNFCCC) provides N_2O emission data from agricultural managed soils reported by countries for a period 1990-2019. Another inventory, the Emission Database for Global Atmospheric Research version 4 (EDGAR v4.3.2), a product of the Joint Research Center and the PBL Netherlands Assessment Agency, contains global N_2O emission inventories (Janssens-Maenhout et al., 2019). EDGAR applies the Intergovernmental Panel on Climate Change (IPCC) guidelines mostly at Tier-1 (using emission factors and activity data). Still, EDGAR integrates higher tier information based on available country reporting, mostly from Annex I countries. EDGAR provides data from 1970 to 2012. The Greenhouse Gas and Air Pollution Interactions and Synergies (GAINS) model (Winiwarter et al., 2018) is an extension of the Regional Air Pollution Information and Simulation (RAINS) model. This extended version (GAINS) estimates emissions of 10 air pollutants and 6 GHGs (including N_2O) for each country based on data from international energy and industrial statistics, emission inventories, and data supplied by countries. GAINS provides N_2O emissions data every five years (i.e., 1990, 1995, 2000, 2005, 2010, 2015) and uses the IPCC emission factor (EF) as a default option for estimation in the absence of more detailed information available for a country.

While there are databases that provide emissions information, each inventory reports N_2O emissions differently. For example, FAOSTAT and UNFCCC provide direct N_2O emissions for the category of synthetic N fertilizers applied to managed soils. In comparison, GAINS and EDGAR report the combined direct N_2O soil emissions from synthetic N fertilizers and manure as fertilizers and crop residues. This difference in emission reporting makes it difficult to compare the estimates of direct N_2O emissions (from synthetic N fertilizers) across different inventories. Additionally, all inventories described incorporate the IPCC Tier 1 guideline for calculating direct N_2O soil emissions, which mainly relies on the self-reported N fertilizer data. Reported data is prone to errors and might be purposely misreported leading to significant uncertainties in estimating N_2O direct emissions in the current approach (Seto et al., 2000). To reduce these uncertainties and provide greater temporal and country coverage, a new approach is needed that considers measurements that are more reflective of synthetic fertilizer usage.

This study introduces a novel methodology for estimating annual, crop-specific N_2O emissions at various spatial resolutions, circumventing the reliance on imprecise self-reported nitrogen fertilizer data. Our approach primarily utilizes crop productivity metrics, offering a more

dependable data source. These metrics reflect the raw materials designated for immediate consumption or integration into broader food systems and supply chains.
The methodology presented has undergone several iterative refinements. Initially, we established a technique for approximating crop-specific nitrogen fertilizer use at a national level. This process employed indicators such as crop yield, cultivated area, and nitrogen use efficiency (NUE). We calculated a country's NUE for a given prediction year based on the prior year's NUE—adjusted using IFASTAT data—and the percentage change in a specific crop's productivity across two consecutive years. Subsequently, we applied the IPCC Tier 1 emission factor to translate nitrogen fertilizer quantities into emission figures. Despite its independence from reported values, this strategy operates at a country-scale, restricting our capacity to identify emission-intensive areas or 'hotspots' within a country. To address this limitation, we devised a strategy for disaggregating national crop-specific nitrogen fertilizer data to a grid scale. This finer resolution was achieved using the detailed crop-specific harvested area data generated by the Global Agro-Ecological Zones (GAEZ) Version 3 model (refer to section 2.2.4).

Nevertheless, the application of the IPCC Tier 1 emission factor continued at this gridded level. To enhance the precision of our emission estimations, we innovated a method that merges experimentally measured data on crop-specific emission factors (EF) with supplementary satellite data from the Google Earth Engine (GEE). This integrated data was then used to train a machine learning model, subsequently employed to estimate emissions at a spatial scale. This comprehensive approach marks a significant advancement in the granularity and accuracy of N_2O emission estimations.

2. Materials and Method

The approach utilized here primarily relies on crop productivity data to estimate N_2O emissions, which is considered a reliable source of information as crop productivity indicates raw materials going for direct consumption or into the food systems and supply chain. We hypothesized that crop productivity could be used as a proxy for estimating N_2O emissions for a country. To investigate this, we selected a subset of high-income, upper-middle-income, lower-middle-income, and low-income countries and classified them by their development level as measured by per capita gross national income (GNI) by the United Nations (UNESCAP, 2020).

To understand the relationship between the average crop productivity and direct N_2O emissions within these selected countries, two variables were used. First, we estimated average crop productivity within the country by accumulating total annual production and harvested area for all the primary crops as defined in FAOSTAT data from 1961-2018 and extracted country-level annual direct N_2O emission data from FAOSTAT for the same period. Then, we compared crop productivity (tons/ha) to direct N_2O emission for each selected country (Figure 1).

Table 1 Pearson correlation coefficients and their significance level between average crop productivity and direct N_2O emission for each country. Correlation coefficients with an asterisk are statistically significant (p<0.01).

GNI level	Country	Correlation coefficient (ϱ)
High-income	Australia	0.69*
	France	0.80*
	Germany	0.18
	U.S.A	0.82*
Upper-middle-income	Argentina	0.92*
	China	0.97*
	Brazil	0.95*
Lower-middle-income	India	0.99*
	Nigeria	0.73*
	Pakistan	0.96*
Low-income	Uganda	-0.09
	Tanzania	0.17

Our analysis found that the average crop productivity and direct N_2O emissions from synthetic fertilizers data were highly correlated for countries in the upper-middle-income and lower-middle-income countries (Table 1). For example, India, China, Argentina, Brazil, and Pakistan have correlation coefficients greater than 0.9. In the high-income countries group, the U.S.A, one of the significant synthetic nitrogen fertilizer users (after China and India), shows a correlation of 0.82 (Figure 1). On the other hand, this relationship was not significant in low-income countries like Uganda and Tanzania. The reason behind this is the minimal use of nitrogen fertilizer in cropland; less than 1% of farmers in most African countries apply fertilizers (Nkonya et al., 2011). One exception was Germany, which belongs to the high-income group and shows no significant relationship between productivity and direct N_2O emissions for reasons unknown to us. In our initial analysis, countries that are significant users of synthetic nitrogen fertilizer, like China, India, U.S.A, Brazil, Pakistan, France, and Australia, which represent 66% of total global nitrogen fertilizer use (FAOSTAT, 2019), show a strong correlation between crop productivity and direct N_2O emissions from synthetic fertilizers. Overall, this analysis supports our hypothesis that crop productivity has a strong correlation with direct N_2O emissions from synthetic fertilizers and it can be used as a proxy to estimate direct N_2O emissions.

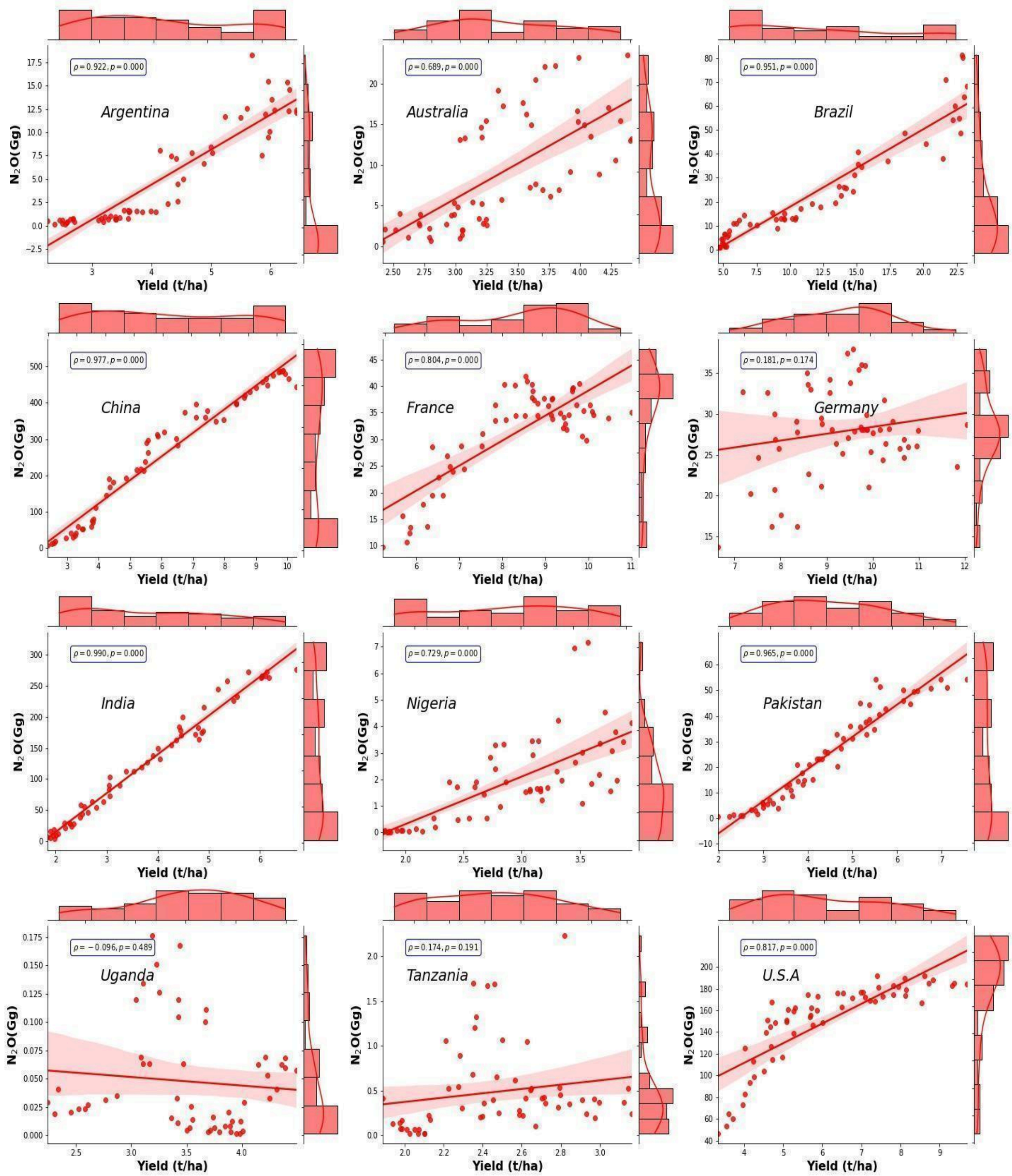

Figure 1 Correlation graph between country-level direct N_2O emissions due to synthetic N fertilizer and crop yield using FAOSTAT data from 1961-2018. (Units- y-axis: gigagram, x-axis: tonnes/ha).

2.1 Data

The provided dataset is essential for creating a model that estimates annual emissions across various spatial scales. Table 2 presents a comprehensive list of datasets employed for the estimation of emissions on a national scale, methodically organized into three distinct segments: model development, calibration, and validation, culminating in the final output. Table 3 details the datasets used for the calculation of annual emissions at a finer gridded scale, approximately 10 kilometers, employing the IPCC Tier 1 emission factor set at 1%. Table 4 outlines the datasets harnessed to derive a spatio-temporally heterogeneous, crop-specific emission factor at an even higher resolution of approximately 4 kilometers.

Table 2 Model development, validation data used for estimation of N_2O emissions at country scale.

Model development	***Unit***	***Source***
Crop-specific yield	Mg/ha,Mg	FAOSTAT
Crop-specific harvest area	ha	FAOSTAT
Crop-specific production	Mg	FAOSTAT, USDA Foreign Agriculture Service
Crop-specific N content	Kg N/ton	Lassaletta *et al.,* 2014
Gridded crop-specific harvest area	ha	Grogan et al., 2022
Model calibration	***Unit***	***Source***
Total N fertilizer at country scale	Kton	IFASTAT
Validation	***Unit***	***Source***
Direct N_2O emission at country scale	Gg	FAOSTAT
Direct N_2O emission at country scale	Kton	UNFCCC
Output	***Unit***	***Source***
Direct N_2O emission at country scale	tCO_2eq /Gg	Model generated

Table 3 Model development, validation data used for estimation of gridded N_2O emissions at ~10 km resolution.

Model development	***Unit***	***Source***
Crop-specific N fertilizer rate at country scale	Kg/ha	This study
Gridded crop-specific harvest area	ha	Grogan et al., 2022
Validation	***Unit***	***Source***

N fertilizer amount at state level USA	Kg	USDA
Output	***Unit***	***Source***
Gridded N fertilizer amount	Kg	Model generated
Gridded direct N_2O emission	tCO_2eq /Gg	Model generated

Table 4 Model development data used for estimation of gridded Emission factor and N_2O emissions at ~4 km resolution.

Model development	***Unit***	***Source***
Global N2O/EF dataset	-	CGIAR
Gridded crop-specific N fertilizer amount	Kg	This study
Crop calendar	date	FAOSTAT
NDVImax	-	Landsat
Annual Precipitation	mm	TerraClimate
Silt	(%)	SoilGrids
Clay	(%)	SoilGrids
Sand	(%)	SoilGrids
pH	-	SoilGrids
Cation Exchange Capacity (CEC)	-	SoilGrids
Topographic Wetness Index (TWI)	-	MERIT Hydro
Output	***Unit***	***Source***
Gridded crop-specific EF	Kg	Model generated
Gridded crop-specific direct N_2O emission	tCO_2eq /Gg	Model generated

2.1.1 FAOSTAT & UNFCCC

FAOSTAT is a data center for FAO (Agency of United Nations), which provides free access to food and agriculture statistics (including crop, livestock, and forestry sub-sectors) for over 245 countries and covers time periods from 1961 to the most recent year available. At the time of this writing, the most recent year for reported synthetic fertilizer emissions per country is 2019. FAOSTAT compiles this dataset using annual questionnaires sent to country focal points within the national statistical systems. The following datasets were used for this study:

a) Crop-specific yield, production, and harvest area data were extracted for primary crops for >100 countries. Primary crops include >150 types; however, 12 nitrogen-fixing (i.e., soybean) crops were excluded from this study as they do not require synthetic N fertilizer application. Instead, these nitrogen-fixing crops obtain N from the atmosphere via the nitrogen fixation process. This information was used in model development (Table 2). For recent years where FAOSTAT crop production data is not updated (results tend to be delayed by 2 years), we use data from USDA Foreign Agriculture Service for global-scale crop production and harvest data. The updated crop data is for the following crops ('Almond', 'Apples', 'Barley', 'Coffee, green', 'Maize (corn)', 'Millet', 'Oats', 'Rice', 'Rye', 'Sorghum', 'Wheat'), which cover a major part of the world crop production.

b) To validate our country-scale emissions, direct N_2O emissions data were extracted from FAOSTAT & UNFCCC for years 2015 to 2019. FAOSTAT calculates direct N_2O emission using Tier 1 approaches, based on generalized emission factors and other parameter values that are specified either globally or regionally. The Tier 1 method employs the default (simplest) method described in the IPCC Guidelines and the default emission factors and other parameters provided by the IPCC. UNFCCC direct N_2O emission data (available via FAOSTAT) was reported by the countries using Tier 2 and 3 methods suggested by IPCC. Both tier estimates were used to validate the modeled approach (Table 2).

2.1.2 IFASTAT

IFASTAT is a data center for the International Fertilizer Association (IFA; https://www.ifastat.org/). This is the only global fertilizer association and has a membership of some 400 entities, encompassing companies across the fertilizer value chain from producers through traders and distributors and service providers to advisors, research organizations and NGOs. Country scale N fertilizer consumption data was extracted from IFASTAT, which is based on a survey conducted every year, sent to country correspondents, including fertilizer associations, fertilizer companies, consultants, experts, and university researchers.

2.1.3 Gridded crop-specific harvest area

Grogen (2022) created a global gridded dataset for the harvested area, production, and yields for 26 distinct crops (Table S2), corresponding to the year 2015. They integrated data from the GAEZ Version 4 global gridded dataset with publicly available information from the FAOSTAT database. This mapping was detailed to a 5-minute resolution and differentiated between irrigated and rainfed crop production systems.

2.1.4 Global experimental N_2O / EF dataset

This database was created as a part of the global N_2O project supported by CGIAR research programs (CRPs) on Climate Change, Agriculture and Food Security (CCAFS). This data compilation effort aimed to create a comprehensive and current repository, capturing extensive field measurements of nitrous oxide (N_2O) emissions from numerous studies that compare fields without fertilizer to those treated with nitrogen (N) fertilizers. This dataset is mined from literature platforms like "ISI-Web of Knowledge," "Google Scholar," and "Scopus," using terms such as "nitrous oxide," "N2O," along with "fertilizer" and "nitrogen usage," and gathered 1153 scholarly articles. Ultimately, after all filtering, this database encapsulates data from 341 research papers, spanning from 1980 to 2016. The database provides information on nitrogen rates, N_2O emissions, and emission factors for various crops across different regions, alongside information on crop types, fertilizer use, and management practices. However, we only use emission factor data and other information, which we will discuss in the next section, extracted from Google Earth Engine (GEE). We constrained the dataset to only for cereal crops (wheat, maize, rice, sorghum, barley, rye). Following is the map of the locations we extracted for our tool development (Figure 2), including all the variables, crops, and years total number of data points were ~4,800.

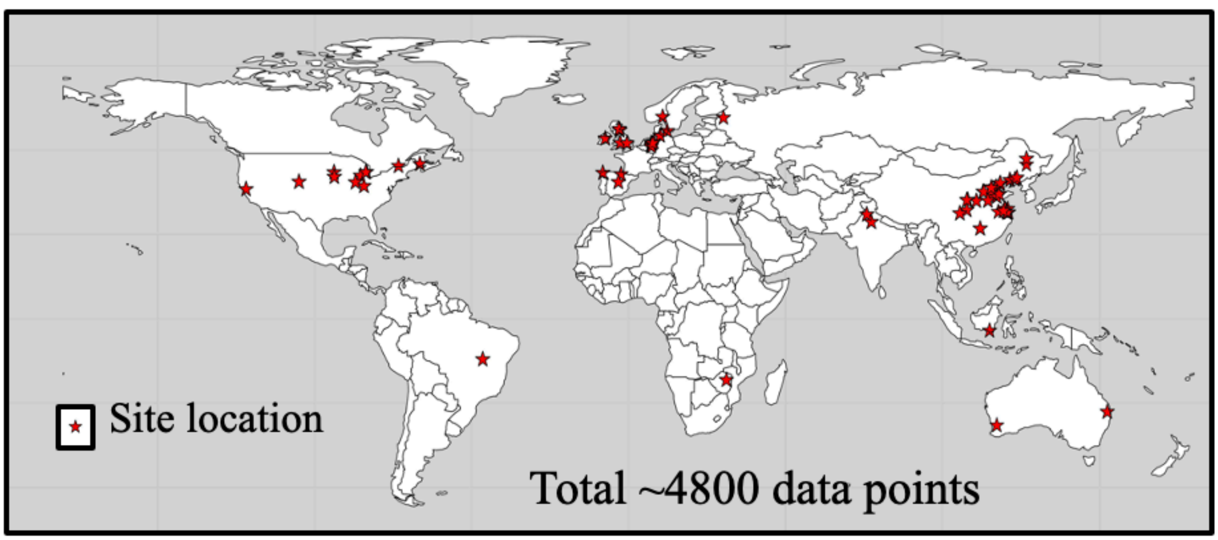

Figure 2 Location for the experimental sites.

2.1.5 Google Earth Engine (GEE)

Google Earth Engine is a cloud-based platform for planetary-scale environmental data analysis. It offers a massive archive of satellite imagery and geospatial datasets with planetary-scale analysis capabilities to detect changes, map trends, and quantify differences on the Earth's surface. The following data we extracted from GEE for our analysis:

2.1.5.1 NDVImax

The NDVImax represents the peak Normalized Difference Vegetation Index (NDVI) value observed during a crop's (corn, wheat and rice) growing season (based on crop calendar. The time series NDVI data was derived from Landsat 8 satellite imagery, utilizing the visual spectral bands. The Landsat 8 mission, a collaborative effort between NASA and the U.S. Geological Survey (USGS), was launched in 2013. This mission is detailed in the work of Marchese et al. (2019). The satellite orbits the Earth with a 16-day revisit cycle and crosses the equator at approximately 10 am, with a possible variation of +/- 15 minutes. Landsat 8 is notable for its two primary instruments: the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). These instruments collectively capture data across 11 spectral bands, ranging from blue to thermal infrared (TIR) wavelengths, approximately 430 nm to 1250 nm. The spatial resolution of the data varies depending on the spectral band, with resolutions of 30m for most bands and 100m for the TIR bands.

For our study, we utilized the Landsat 8 Collection 2, Tier 1 Top of Atmosphere (TOA) data, which was accessed through the Google Earth Engine (GEE). Further information about the Landsat 8 mission and its data products is available on the USGS website at USGS Landsat 8. More information and application of NDVImax is in section 2.2.6.

2.1.5.2 Soil information

Soil information, including soil texture, pH, and CEC, was derived from SoilGrids dataset. SoilGrids is a cutting-edge system dedicated to digital soil mapping on a global scale. It employs advanced machine learning techniques to create predictive models of soil properties distributed around the world. These maps detail soil characteristics at six standard depth intervals, which adhere to the specifications outlined by the GlobalSoilMap project of the International Union of Soil Sciences (IUSS) Working Group. The maps are rendered with a fine spatial resolution of 250 meters, providing detailed and actionable insights into soil composition and structure for various applications. More information about the data can be find on the following link : https://www.isric.org/explore/soilgrids/

2.1.5.3 Climate data & TWI

Climate data which include annual precipitation is derived from TerraClimate. TerraClimate is a collection of data that provides monthly updates on climate conditions and water balance for land areas around the world. It improves accuracy by blending detailed climate patterns from the WorldClim dataset with broader, time-changing data from two other sources: CRU Ts4.0 and the Japanese 55-year Reanalysis (JRA55). More information about the data can be find on the following link : https://www.climatologylab.org/terraclimate.html

TWI is calculated based on flow accumulation area and slopes obtained from a global hydrography dataset called MERIT Hydro. MERIT Hydro presents a novel global flow direction map with a fine resolution of approximately 90 meters at the equator, equivalent to 3 arc-seconds. This map is crafted using the most recent topographical data from MERIT DEM, alongside comprehensive water body datasets which include G1WBM, GSWO, and data from

OpenStreetMap. More information about the data can be find out on the following link: https://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_Hydro/index.html

2.2 Method

This novel method is based on an integrated approach using country-level crop harvest area, yields, and crop-specific NUE, to predict the amount of synthetic N fertilizer applied to the primary crops in a country for a particular year. Country-level estimates of N_2O emissions are based on the IPCC (2006) emission factors for synthetic N fertilizer use (Figure 3). Using this approach, we estimated the emissions for 103 countries which represent approximately 98% of total synthetic N fertilizer use globally based on 2019 FAOSTAT data.

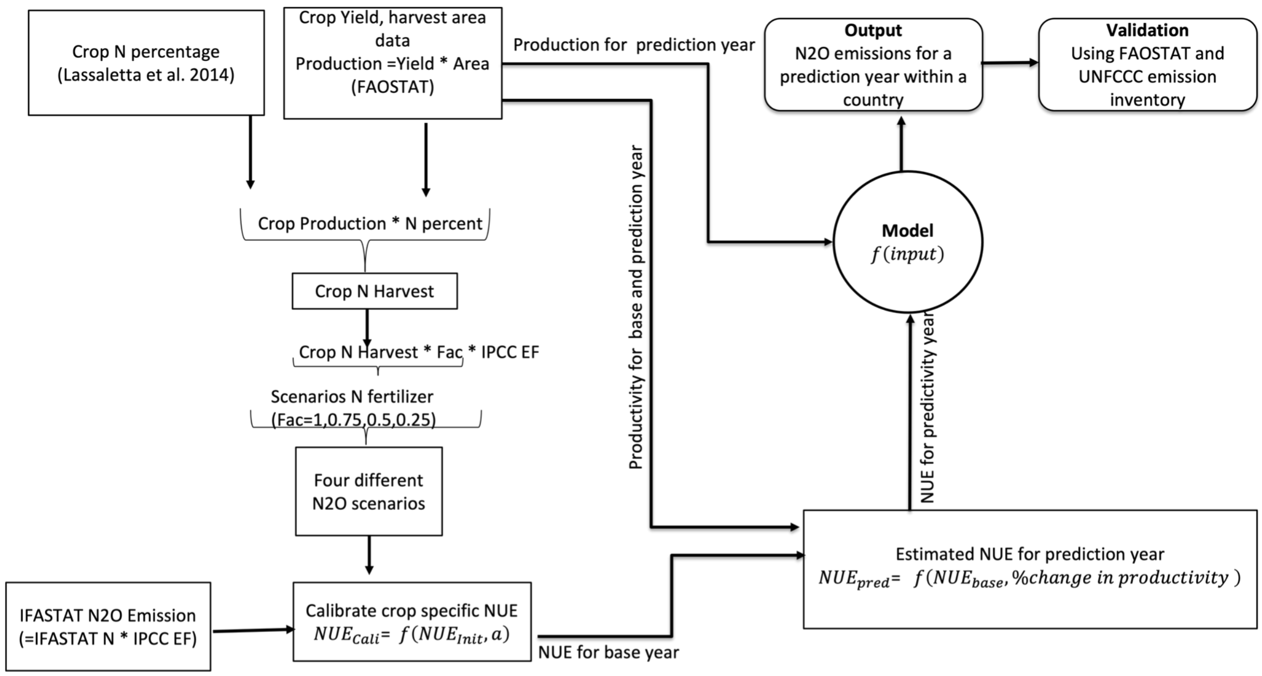

Figure 3 Schematic representation of estimating N_2O emissions from crop productivity at country scale.

2.2.1 Scenario development for N fertilizer amount

The first step in this method was to estimate N fertilizer applied in a country. To do this, four scenarios were created to estimate N fertilizer applied as the fraction of N uptake into crops. To develop each scenario, first, we used FAOSTAT crop-specific yield and harvest area data, and nitrogen content ($N_{CONTENT}$; Eq. 1; Table 2) to calculate the amount of N removed from the field during the harvest process, denoted as $N_{HARVEST}$. For this calculation, it was assumed that $N_{HARVEST}$ is the actual N uptake from the soil ($\sim N_{UPTAKE}$). N taken up in crop residues and roots were excluded from this accounting of N uptake, as this N is assumed to be returned to the soil, thus cycled in, a common accounting procedure in N mass balance (Basso et al., 2019). The total N uptake in the crops during the growing season can be from many N sources like Synthetic N fertilizers, Manure N, N deposit, and N mineralization from the soil. However, Synthetic N fertilizers are recognized as the most important factor contributing to direct N_2O emissions from agricultural soils. Based on our assumption that the $N_{HARVEST}$ is proportional to the amount of synthetic N fertilizers applied, four scenarios were designed for N fertilizer (N_{FERT}) as shown in Figure 4.

$$N_{FERT} = \underbrace{\text{Yield} \ \text{x} \ \text{Area} \ \text{x} \ N_{CONTENT}}_{N_{HARVEST} \ (\sim N_{UPTAKE})} \ \text{x} \ \text{Fac (Eq. 1)}$$

Where, "Fac" is the proportionality factor for each scenario (1, 0.75, 0.5, 0.25). Then each scenario is converted into emissions using the emission factor (EF) from IPCC (2006) in Eq. 2:

$$N_2O = N_{FERT} \ \text{x EF (Eq. 2)}$$

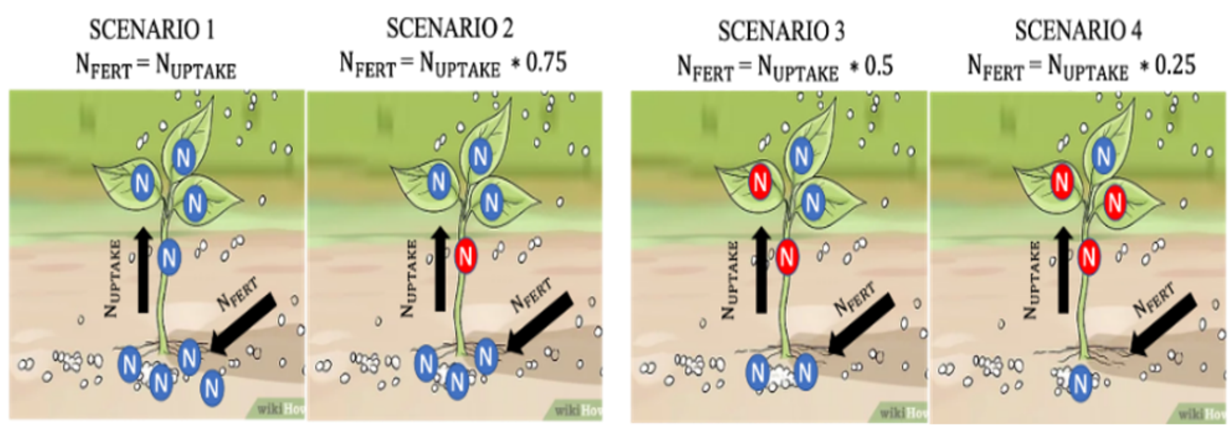

Figure 4 Overview of N fertilizer amount as a proportion of N uptake for four scenarios. For Scenarios 1, 2, 3, 4, the N fertilizer amount is 100%, 75%, 50%, and 25% of N uptake by the crops, respectively.

Scenario 1 represents the ideal case where all the N taken up by the plants is supplied from the synthetic N fertilizer applied and both are equal in amount. In scenario 2, 3, and 4 we considered 75%, 50%, and 25%, of the N taken up by the plants is supplied from the synthetic N fertilizer applied, respectively. For example, suppose a case where actual emissions (calculated based on

IFASTAT fertilizer data) in three countries are higher than, equal to, and less than scenario 1. Then, in the first country, where emissions are greater than scenario 1, it suggests that the N fertilizer amount applied is more than the required amount needed by crops for their growth. But, on the other hand, the second country, where emissions are equal (or closer) to scenario 1 (ideal case), represents the adequate application of fertilizers, matching the amount needed for crop growth. For the third country, where the emissions are less than scenario 1, this suggests a lower fertilizer application rate than the crops require. We further divided the last case into three more scenarios 2, 3, and 4, which represent the case where the amount of fertilizer applied is 25% (75% of N uptake is from synthetic fertilizer), 50% (50% of N uptake is from synthetic fertilizer), and 75% (25% of N uptake is from synthetic fertilizer), less than the required amount, respectively.

Our aim in creating these "Scenarios" is to understand how the actual emissions data compares to our assumed scenarios, and it might or might not represent the real-time conditions. These scenarios were compared to real-time emissions to see how well each scenario relates, which provides insight into the N fertilizer application amount to the crops within a country. This can help determine if the amount applied is higher, equal, or lower from the required amount needed for the crops.

2.2.2 Update IFASTAT total N fertilizer data excluding grass application

Our model calibration uses the total nitrogen (N) fertilizer amounts from IFASTAT, which include the quantities applied to grasslands. However, our primary focus is on the N fertilizer amounts used on croplands. This approach can lead to overestimation in countries like the Netherlands and New Zealand, where grasslands consume a significant share of the total N fertilizer. To address this potential overestimation, we adjust the total N fertilizer figures by subtracting the portion used for grasslands. For calculating this grassland fraction, we rely on IFASTAT data that provides details on the total N consumption in a country, including the specific amounts used on grasslands. In cases where IFASTAT does not have data for a country, we use alternative data from Lassaletta et al., 2014, which offers information on the fraction of total N fertilizer applied, drawing from various compiled sources.

2.2.3 Estimation of crop-specific NUE at a country scale

This step provides insight into how efficiently each primary crop takes up the amount of N fertilizer applied within the country and its temporal variability. Country-scale nitrogen use efficiency (NUE) was calculated using N fertilizer amount in scenario 1 which is equal to N uptake for each crop (Fac =1; Eq 1) and observed N fertilizer from IFASTAT (Eq. 3):

$$\text{NUE} = \frac{N_{FERT}\,(Scenario\ 1)}{N_{FERT}(IFASTAT)} \qquad \text{(Eq. 3)}$$

Our calculated NUE is most relevant to Synthetic N fertilizers, ignoring other N inputs (i.e., N manure, deposits). However, as our target was to estimate direct N_2O from Synthetic N fertilizers, it is relevant to our final calculations.

Based on the understanding acquired about the country-specific NUE, we scaled up or down the crop-specific N fertilizer amount from scenario 1 to a new value. Also, we make sure that the scaling factor for the individual crop is weighed by its contribution to total N_2O emissions in that scenario. Using this approach, we calculated the NUE for each crop within a country.

Our aim here was to develop an independent estimate of N fertilizer amount for the estimation of N_2O emissions. Therefore, a country's predicted NUE for a given year can be explained by the changes in crop yield from base year to prediction year.

2.2.4 Prediction of N_2O emissions

In the final step, we estimated annual emissions for a country. First, we calculated the emission for each crop using yield, harvest area, and NUE (Eq. 4). Then, we summed the crop-specific N_2O for all the primary crops within the country to determine the total emission (Eq. 5). To estimate direct N_2O emissions for a country, the following input data for each crop was needed:

1. Yield, harvest area, Nitrogen content for target year (Table 2)
2. Estimated crop-specific NUE for the target year (estimated in section 2.2.2)

The inputs from 1 and 2 above were combined in Eq.4 to estimate N_2O emissions:

$$N_2O_{i,j} = \left(\frac{Yield_{i,j}\ x\ Area_{i,j}\ x\ Ncontent_i}{NUE_i}\right)\ \text{X EF (Eq. 4)}$$

Where EF converts N to N_2O. The final, total, calculation for a country,

$$TotO_j = \sum_{i=1}^{n} N_2O_{ij}\ \text{(Eq. 5)}$$

Where i is the type of crop and j is the target year for the country. N_2O_{ij} is the direct emission predicted for the particular crop (i) in a target year (j), and n is the total number of primary crops defined in FAOSTAT data within a country. $TotN_2O_j$ is the total direct emission for a country in a target year (j).

2.2.5 Downscaling from country scale to grid-level

Once country-level emissions estimates were generated, we downscaled the country-level N fertilizer amount to grid level at a spatial resolution of 0.083° (~10 km) in latitude by longitude, using country scale crop-specific N fertilizer rate (calculated in earlier steps) and gridded harvest area produced by Global Agro-Ecological Zones (GAEZ) Version 3 model, which is based on the FAO crop production data (Figure 5).

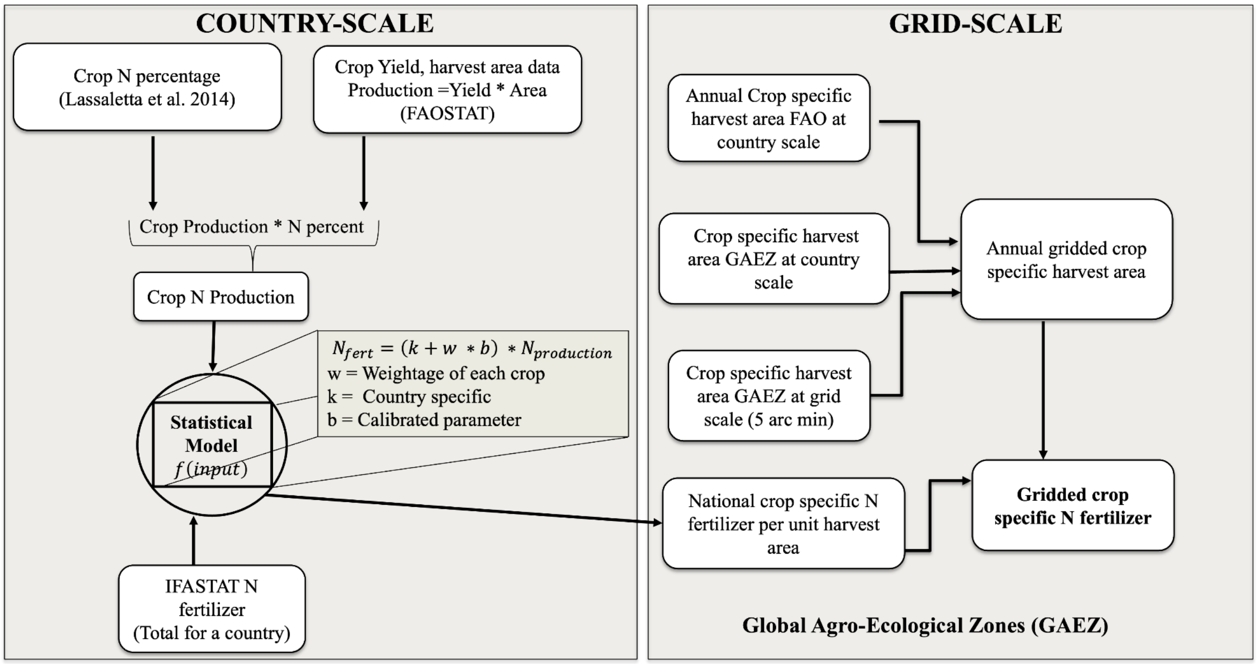

Figure 5 Schematic representation of estimating N_2O emissions from country-scale to asset level.

GAEZ crop classification is grouped into 26 crop-types which includes all the 160 unique crops from FAOSTAT (Table S3). This was done to understand the spatial distribution of N_2O emissions within a country and to identify areas with significant N_2O emissions (fertilizer application). Our study has excluded the nitrogen fixing crops (discussed in section 2.1.1) from FAOSTAT data. Therefore, to harmonize our crop type to GAEZ classification we distributed our crop type into 23 crop groups from GAEZ data, excluding soybeans and pulses. Then, we aggregated country scale crop-specific N fertilizer data (estimated in our study) and harvest area (from FAOSTAT) for each group of crops. Next, using these two data sets we calculated the country scale N fertilizer rate for each crop group. Gridded harvest area for 2015 to 2020 and crop-specific FAOSTAT country scale harvest areas were used to upscale or downscale the gridded values. Then, the crop-specific gridded harvest area was multiplied with N fertilizer rate (country-specific) to estimate crop-specific N fertilizer amount at 0.083° (~10 km spatial resolution) and aggregated for all the crops within the grid to derive the total N fertilizers applied. Finally, to calculate direct N_2O emission from Synthetic N fertilizers, the aggregated N fertilizer amount within the grid was multiplied with the IPCC emission factor.

2.2.6 Spatio-temporal heterogeneous crop-specific EF

Until this point, we have been using static IPCC tier-1 EF 1% to calculate emissions from N fertilizer, which doesn't vary spatially or temporarily. However, EF value 1% is recommended by IPCC to calculate emissions where tier 2 and 3 information is not available. To further refine our methodology, we developed a tier 3 method (Figure 6) which combined N_2O monitoring from experimental sites, alongside remote-sensing data for climate, soil, topography and vegetation growth index (discussed in section 2.1), to train a validate a machine learning model to predict crop-specific emission factors.

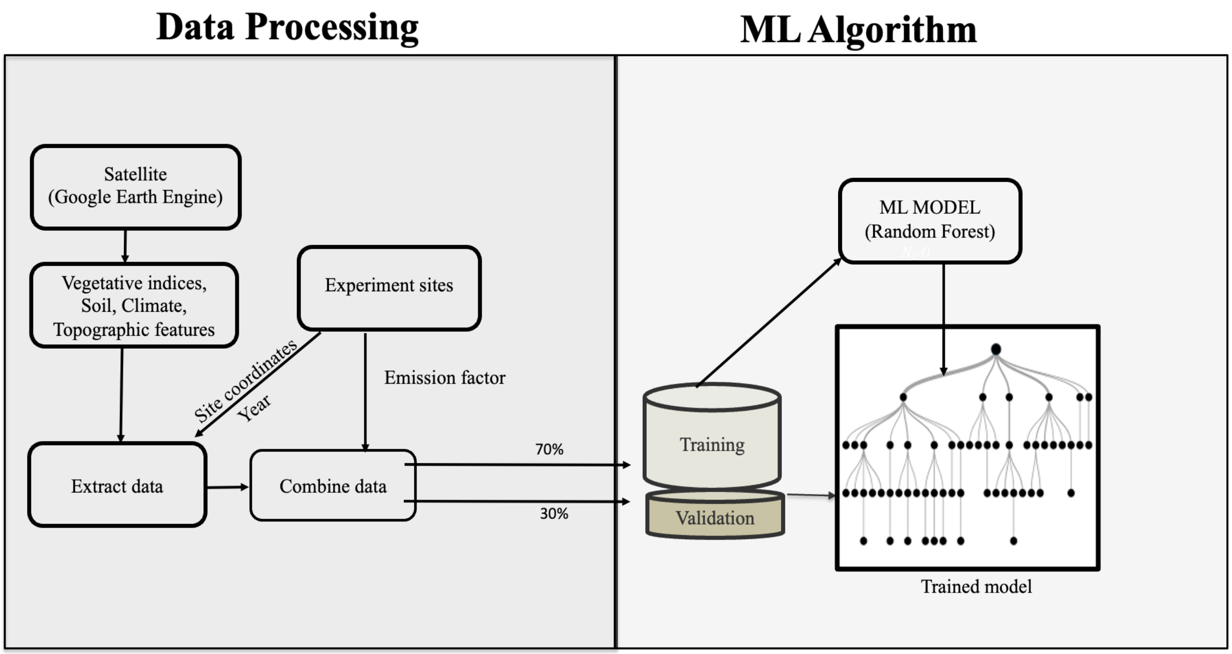

Figure 6 Schematic representation of estimating crop-specific EF at gridded level.

2.2.6.1 Data processing

Data was gathered from various sources, as detailed in section 2.1, with each dataset having its own spatial resolution. Our initial step involved calculating the maximum NDVI (NDVImax) for each crop's growing season using Landsat 8 imagery (section 2.1.5.1**)**. The FAO crop calendar provided the necessary planting and harvesting dates for each country. With these dates, we refined the NDVI images to capture only the relevant season, and from these, we determined the peak NDVI value for each grid. This process allowed us to generate a country-specific map showing the NDVImax for each crop. Subsequently, for sites where crop-specific EF was measured, we pinpointed the exact latitude and longitude and gathered all pertinent data, including soil characteristics, climate, topography, and NDVImax. This comprehensive dataset was then prepared and organized, making it suitable for analysis with machine learning models.

2.2.6.2 ML model training and testing

A Random Forest (RF) algorithm was utilized to predict crop-specific EFs. The Random Forest regression model is a robust and versatile machine learning algorithm that operates by constructing a multitude of decision trees at training time and outputting the mean prediction of the individual trees. This ensemble approach leverages the power of multiple decision-making units, thereby reducing the risk of overfitting and enhancing the predictive accuracy of the model. Random Forest regression is particularly adept at handling large datasets with numerous features, as it can capture complex, non-linear relationships. It also provides useful insights into feature importance, which can be valuable for understanding the driving factors behind the model's predictions. In our first step we distributed the data into training (70%) and testing (30%). We also performed hyperparameter tuning using 5-fold cross validation to ensure against over-fitting. Figure 7 shows the ML model performance in predicting EF with testing data with significant correlation of 0.61 and root mean square error (RMSE) of 0.42 with observed data. The right plot shows the feature importance. Our initial hypothesis for using NDVImax as one of the features in the ML model was that it can be an indirect proxy of available N in the soil. For example, if a crop has higher NDVImax for a season, it may indicate good crop health and higher biomass accumulation, indicators that a crop will more efficiently take up nitrogen. This intuition appears to be confirmed, as the ML model determined NDVImax was the most important feature in terms of explaining the variability of the EF. The next important feature is CEC, which is a measure of how many cations (positively charged ions) can be retained on soil particle surfaces. Negatively charged sites on the surfaces of soil particles, such as clays and organic matter, hold onto cations like calcium (Ca^{2+}), magnesium (Mg^{2+}), potassium (K^+), and ammonium (NH_4^+). CEC can be seen as a factor that mediates many soil processes, including those that lead to N_2O emissions. Soils with high CEC might help to mitigate N_2O emissions by retaining ammonium and stabilizing pH, but this can vary widely depending on soil type, climate, plant cover, and management practices. Similarly, annual precipitation is quite important in terms of N_2O emissions. It can lead to increased soil moisture, which displaces oxygen and creates anoxic conditions that are favorable for denitrification which is a major contributor of N_2O emissions. Soil pH affects N_2O emissions by influencing the microbial processes of nitrification and denitrification, the availability of substrates for these processes, enzyme activity, and the competition between different microbial groups. While these features were chosen carefully, the relationship between soil properties such as pH, moisture content, cation exchange capacity (CEC), climate, and other factors with N_2O emissions is indeed complex and often non-linear. Finely tuned ML models, such as the Random Forest Model we developed, are well suited to capture these kinds of complex, non-linear relationships.

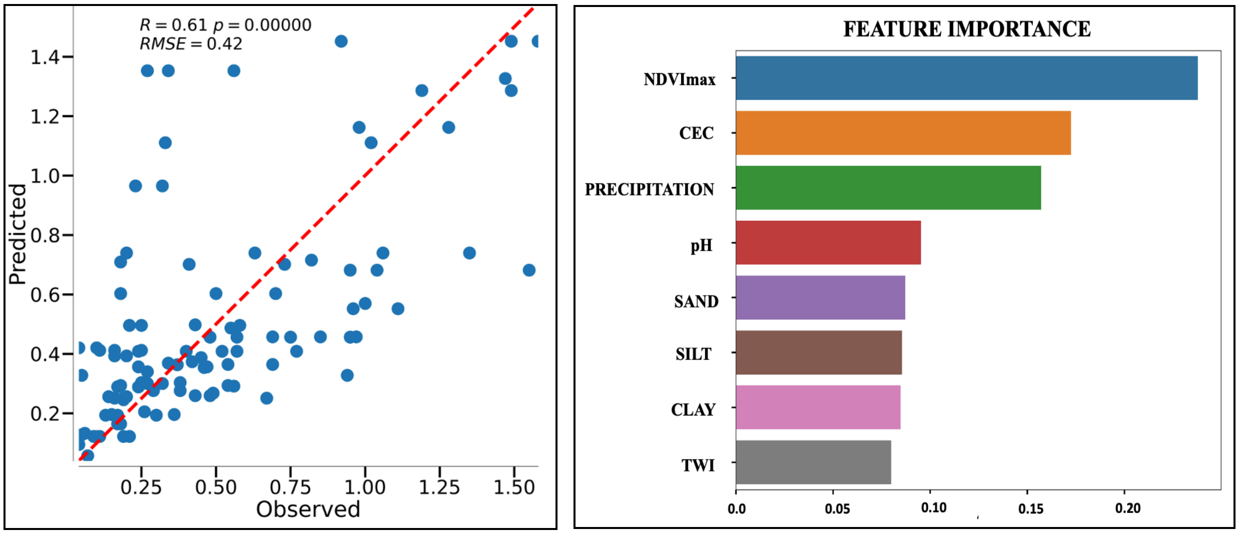

Figure 7 ML model performance with testing dataset (left) and feature importance plot (right) suggesting the dominant drivers behind the model prediction.

2.2.6.3 Scale up the model to create EF map

After training and testing the machine learning model, we applied the model on a spatial scale to generate emission factor (EF) maps. Figure 8 displays the EF maps for corn, wheat, and rice across the countries of the USA, India, and China for the year 2021. These maps reveal the considerable spatial heterogeneity of EFs within each country, highlighting the variability of emissions even within the same crop type. The maps underscore the importance of region-specific analyses for a more accurate assessment of agricultural emissions and underscore the need for nuanced EFs that can be applied to diverse agricultural practices, soil types, and climatic conditions.

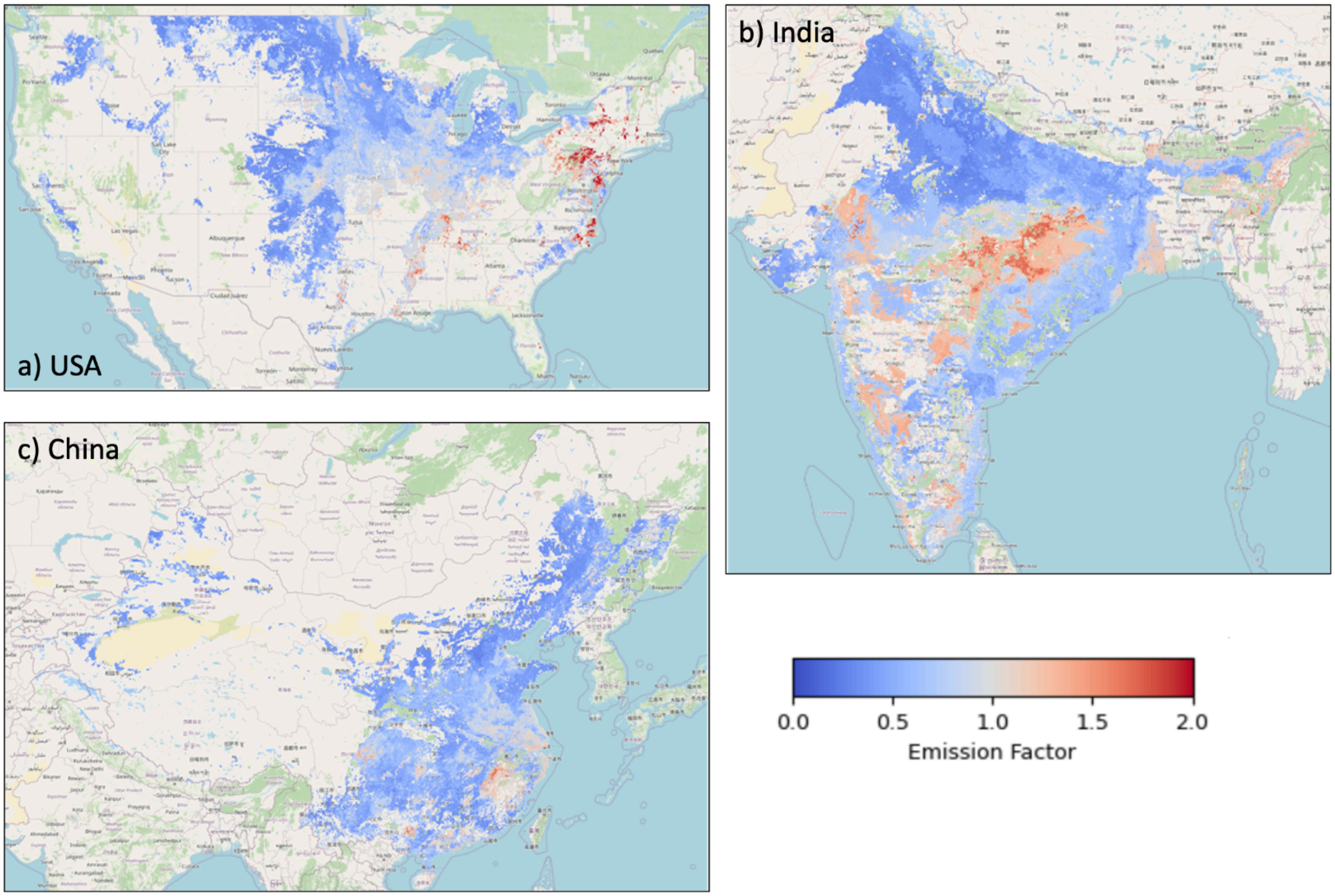

Figure 8 EF factor map for corn, wheat and rice with countries a) USA, b) India c), China for 2021.

2.4 Verifying modeled emissions estimates

To verify the modeled country-level N_2O emission estimates, FAOSTAT and UNFCCC direct N_2O emissions data were used. A time-series analysis was performed for 24 countries, representing the significant portion of N_2O emissions globally, and a direct modeled emissions to FAOSTAT or UNFCCC comparison was performed (Figure 12). In addition, for all the countries (where data were available) one to one comparison is performed with FAOSTAT and UNFCCC (Figure 6; discussed in detail in next section). The spatially gridded direct N_2O emissions were not compared to other emissions estimates (i.e., *in-situ* measurements) due to a lack of globally gridded distributed data at the time of this work. However, we used USDA crop-specific state level N fertilizer data in the USA to validate our estimated N fertilizer amount aggregated at state level. Our estimates from corn and wheat are compared for 25 states in the USA including years from 2015-2018 (Figure S2).

3. Results & Discussion

3.1 Country scale

We compared N_2O derived from the N_{FERT} amounts in different scenarios to the IFASTAT observation data at the country-level (Figure 11a). IFASTAT provides N fertilizer consumption data at the country scale. We converted fertilizer consumption data to N_2O emissions by applying the IPCC (2006) emission factor. Figure 9b, shows the country scale NUE (calculated in section 2.1.2). For Asian countries, like India, China, Pakistan, and Indonesia, the direct emissions (or N fertilizer) observed is much higher than scenario 1, suggesting primary crops take up a very low percentage of N fertilizer applied (low Nitrogen Use Efficiency: NUE<<1).

On the other hand, direct emissions for Tanzania, Kenya, and Nigeria, are closer to scenario 4 for most years, suggesting a small fraction of N uptake comes from synthetic N fertilizer (lower N application) in these countries (N mining: NUE >1). Soil N mining or fertility loss results from over exploitation of agricultural land without giving back the lost N to the soil. Limited access of synthetic N fertilizer and increasing population pressure led to severe loss of soil nutrient fertility in these countries. This problem of nutrient mining and loss of soil fertility in many African countries is highlighted in previous studies (Vitousek et al., 2009, Liu et al., 2010).

In countries like the U.S.A, France, Canada, Germany, Australia, Italy, and South Africa, observed emissions are closer to scenario 1, suggesting these countries have higher nutrient efficiency (Moderate Nitrogen Use Efficiency: NUE < 1) as compared to Asian countries. One of the reasons might be due to the use of better nutrient management practices in these countries.

According to our analysis, in the last two decades, NUE for most countries have not varied significantly, except for countries like Nigeria and Tanzania, where synthetic fertilizer is not the primary source of crop N uptake, and extra N comes from the agricultural soil nutrient storage. Lassaletta et al. (2014) showed the NUE trends for the past 50 years for 124 countries and found higher crop yield than fertilization (NUE >1; N mining) for 18 countries such as Canada, Morocco, Algeria, Iraq, and Mozambique in the 1960–1980 period (Lassaletta et al., 2014). Also, in recent years of his study period, NUE higher than 1 have been observed in 10 African countries and former Soviet Union countries, Afghanistan and Paraguay. Argentina is the only country with NUE higher than 1 for the whole period of his study. Our study and results shown in Figure 9b suggests NUE is greater than 1 for countries like 1) Argentina, 2) Nigeria and Tanzania (African countries), 3) Russia and Ukraine (former Soviet Union countries), aligning well with analysis from Lassaletta et al. (2014). Note that our calculated NUE ignores other N inputs beyond synthetic fertilizer (i.e., N manure, deposits), which could influence our assessment of NUE in systems that rely on organic amendments.

a)

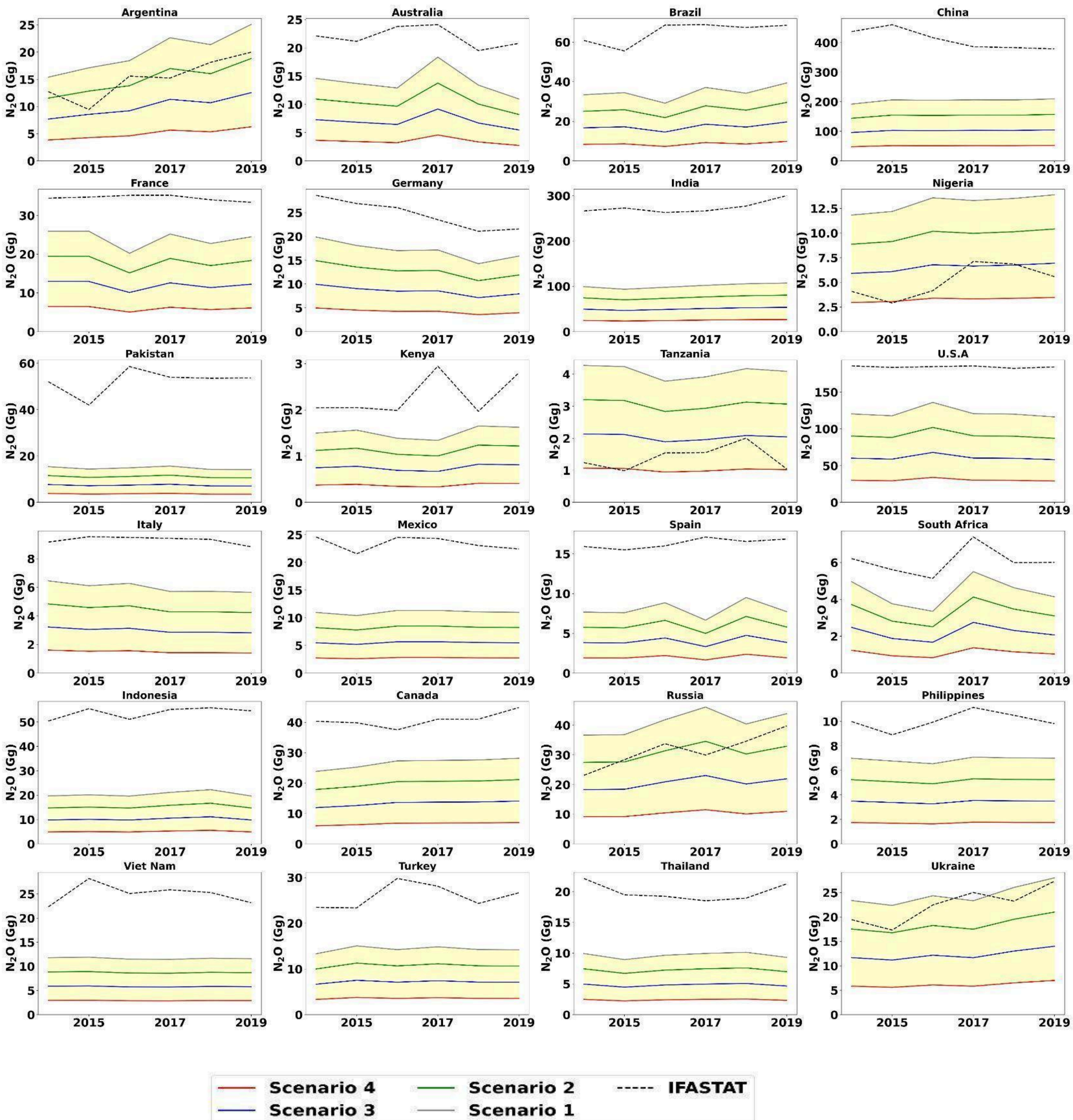

Figure 11 cont.

b)

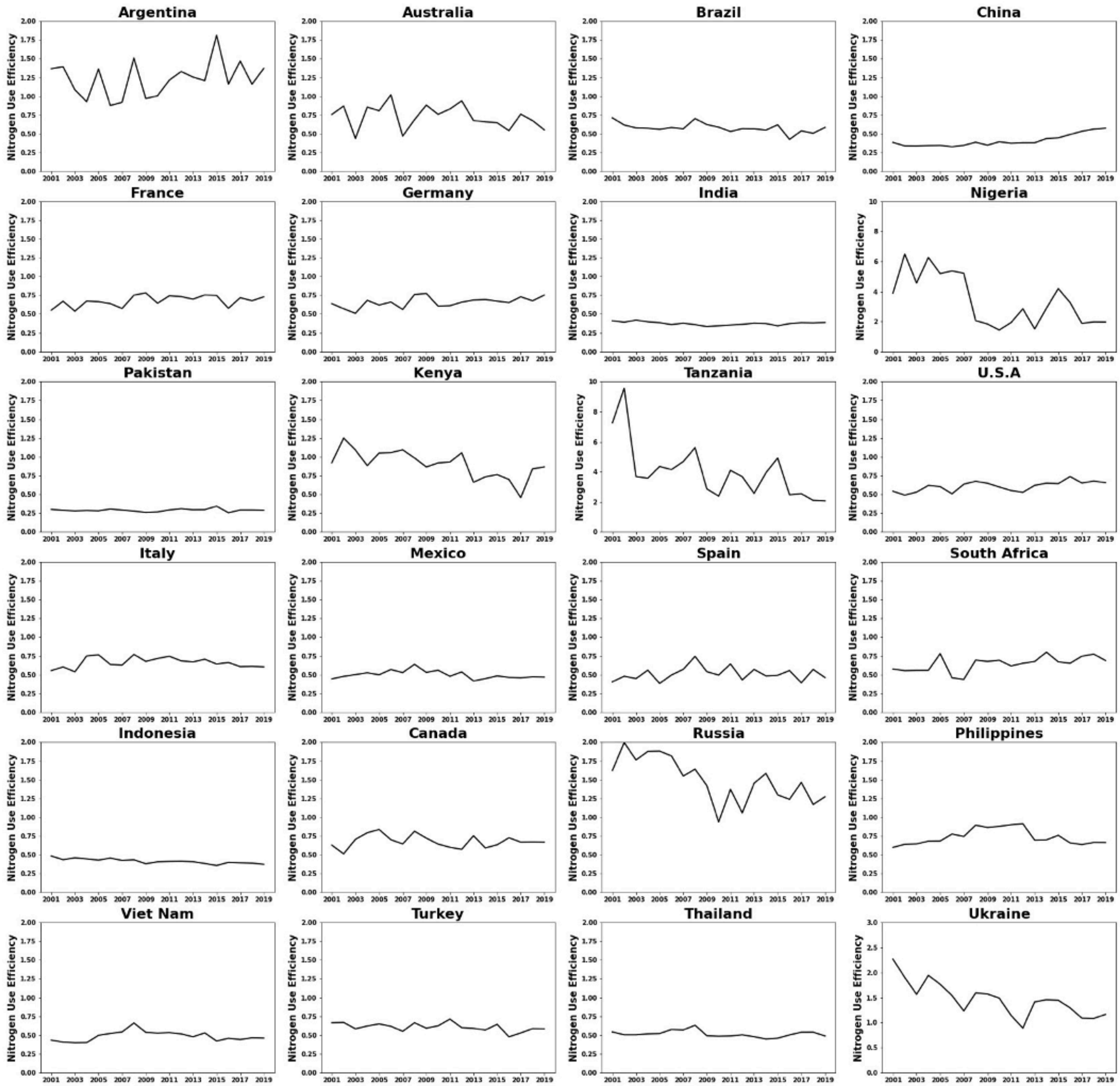

Figure 9 a) Time-series (2014-2019) of N_2O from IFASTAT and N_2O estimated from scenarios 1, 2, 3 and 4. Note, y-scales, N_2O (Gg) are different for each country; **b)** Nitrogen use efficiency (NUE) for each country based on scenario 1 NFERT and IFASTAT. Note, y-scales are different for each country.

In this study, we include 103 countries, which represent the majority of total synthetic nitrogen fertilizer use at global scale (~99%). According to our estimation, total direct N_2O emissions from synthetic nitrogen fertilizer use combined for 103 countries are 1,623, 1,610, 1,636, 1,585, 1,600, and 1,652 gigagrams (Gg) for years 2015, 2016, 2017, 2018, 2019, and 2020 respectively. China, India, and the U.S.A. are the top three emitter countries for the period 2015 to 2020 contributing more than 50% of total emissions every year (Table 6). Including the next set of highest emitters - Brazil, Pakistan, Indonesia, and Canada - the total percentage contribution reaches more than 60% of total emissions (>900 Gg) every year (Table 7).

Table 7 List of the 7 countries which contributed at least 69% to the total global direct N_2O emissions due to synthetic nitrogen fertilizers during the period of 2015-2020. Values are measured as percent (country total/global total *100).

Country	Percent contribution to total emission by year					
	2015 (%)	**2016 (%)**	**2017 (%)**	**2018 (%)**	**2019 (%)**	**2020 (%)**
China	29.1	28.2	25.3	24.0	23.6	22.8
India	16.3	16.7	16.3	16.6	17.3	19.0
U.S.A	11.2	11.9	10.6	11.6	11.3	11.2
Brazil	3.7	3.3	4.7	4.1	4.5	4.3
Pakistan	3.2	2.6	3.5	3.3	3.3	3.3
Indonesia	3.1	3.3	3.4	3.6	3.1	3.2
Canada	2.5	2.4	2.4	2.7	2.5	2.7

In addition to country-level estimation, the accumulated emissions per year at a continent level for Asia (AS), Europe (EU), America (AM), Africa (AF), and Oceania (OC) are provided in Table S1. AS is the highest emitting continent with a total emission of 949 Gg or 58.7% of total global emissions(averaged over the study period). AM, EU, AF, and OC emit 360 Gg (22.3%), 233 Gg (14.5%), 51 Gg (3.2%), and 21 Gg (1.3%), respectively (Figure 10). Even though AF is quite significant in the land area, it only represents 3.2 % of total global emissions due to the limited access to synthetic use N fertilizers in African countries. AS and AM continents collectively represent more than 80% of the total global emissions.

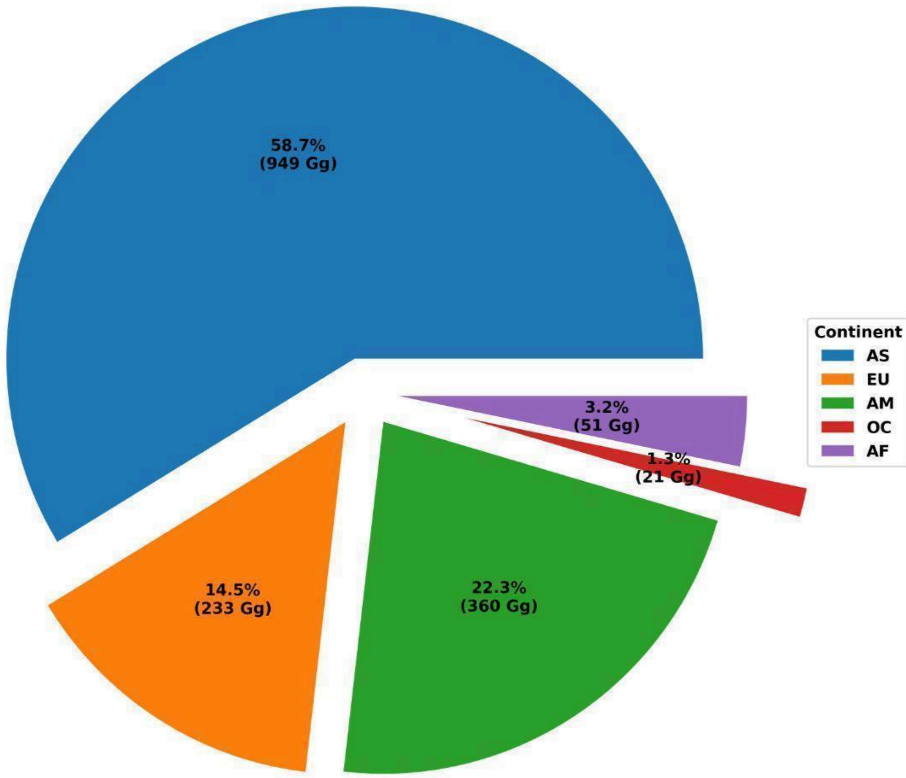

Figure 10 Average percentage contribution of each continent to global total direct N_2O emissions averaged over the period of 2015-2020. In parentheses are the average amount of emissions (Gg) for the same period. Legend key as follows: Asia (AS), Europe (EU), America (AM), Africa (AF), and Oceania (OC).

Our predicted N_2O emissions are independently compared against country-level estimates provided by FAO and UNFCCC inventories (Figure 11). FAO and UNFCCC provide direct N_2O emissions estimation from synthetic nitrogen fertilizer use at country scale. FAO estimates are computed at Tier 1 following the 2006 Guidelines for National GHG Inventories of the Intergovernmental Panel on Climate Change (IPCC, 2006). The primary data source for FAO is national data collected via the FAO Fertilizers questionnaire sent to FAO focal points in the national government. The UNFCCC provides direct N_2O emissions from managed agricultural soils reported by the countries. In this inventory, Annex I group countries' direct N_2O emissions from agricultural soil are calculated by combining IPCC's higher tier approaches (Tier 2 and 3). UNFCCC has limited countries which separately report direct emissions for the category of inorganic N fertilizers applied to managed soil.

Overall, the model estimated emissions correlate with FAO and UNFCCC data inventory with an R^2 of 0.94 and 0.96, respectively ($p<0.001$). As discussed earlier, according to our estimations, China, India, U.S.A, and Brazil are the top four emitters globally (Table 6). When we compared our estimation with the FAO database, we are slightly underestimating the emissions for a few countries. However, our estimates are close to UNFCCC for the other two countries, India and U.S.A. For countries like the U.S.A and Australia, our estimates are more comparable to FAO than UNFCCC, which might be because we are following the IPCC Tier 1 approach to calculate the final direct emissions from N fertilizer amounts like FAO. UNFCCC reporting countries use IPCC's higher tier approaches (Tier 2 and 3) for emission estimates.

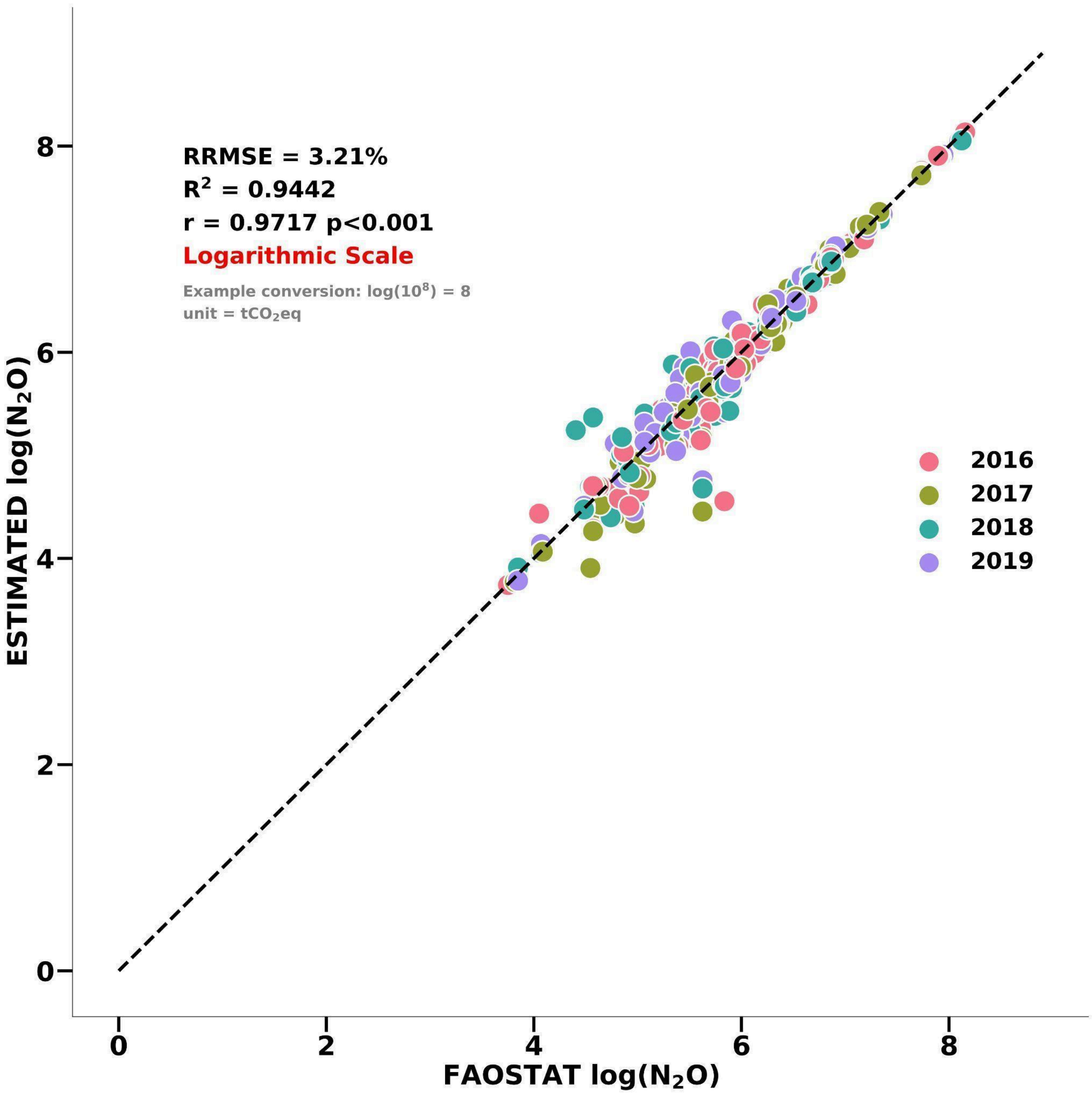

Figure 6 cont.

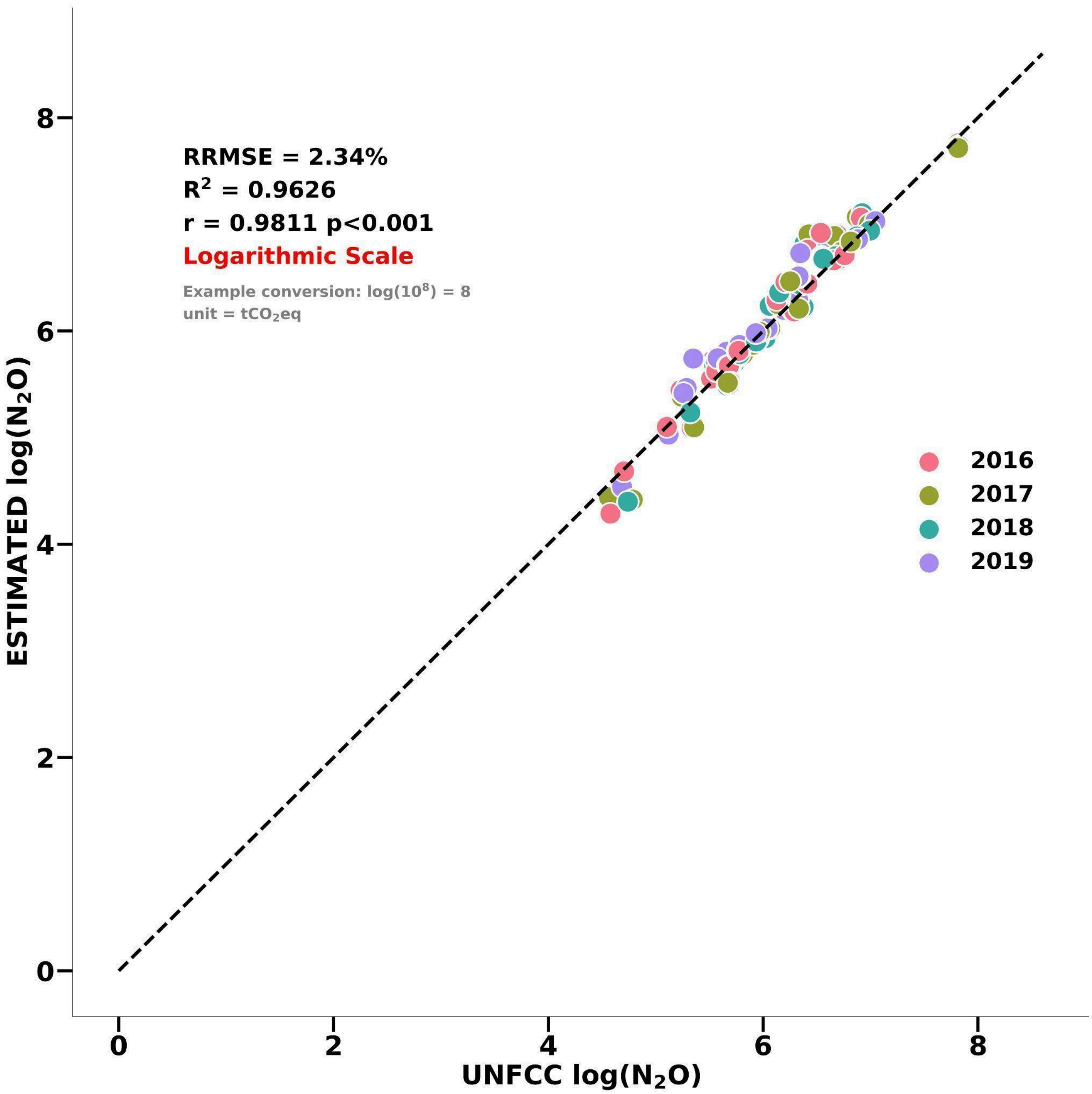

Figure 11 The model predicted N_2O validation with FAOSTAT (top image) and UNFCCC (bottom image) at the country-level for the time period 2016-2019. Dashed black line is the 1:1: line, indicating perfect agreement. Note, each figure has different scales.

To further analyze the results, we compared a time series of emissions for 24 countries from 2015 to 2019 with FAO and UNFCCC inventory (Figure 12). These 24 countries are selected from each of the five continents (Asia, America, Europe, Africa, and Oceania). According to our calculation, in total, these 24 countries represented 84%-86.4%, total global emissions between 2015 and 2020.

The discrepancies between our estimates and UNFCCC might be due to the difference in method. For example, countries like the USA use different approaches for national scale N_2O emissions assessments, which involve process-based models, such as the DAYCENT (Del Grosso et al., 2012), along with the IPCC emission factor. For some countries UNFCCC emission values are not available between 2015 to 2019, and we compared our emission with FAOSTAT only. In China, which is the highest emitting country (as discussed earlier), our estimates underestimate emissions as compared to FAOSTAT. However, both are showing decreases in emission levels from 2015 to 2019. Other high emitting countries like India and Pakistan show a close match to FAOSTAT estimates.

In addition to time-series comparison, to further analyze the changes in emissions in the last five years, we have compared the percentage change in emissions between five years from 2015 to 2019 to FAO and UNFCCC (if available; Table S2). Positive or negative percent change indicates increase or decrease in emissions level between 2015 and 2019, respectively. According to our estimates 14 countries out of 24 are showing the same sign (positive or negative) of percent change in emissions compared with FAO or UNFCCC. Another contrasting difference between our method and FAO estimates is the temporal variability in the emissions. Our model suggests high temporal variability in emissions for countries like Argentina, Australia, and Brazil compared to FAO estimates. The reason could be that our predicted NUE for crops depends on the change in crop yield from one year to another. Therefore, if yield changes from one year to another within a country, as it normally does, changes in the emissions are reflected in our estimates. To further validate our estimations, we compared crop-specific country-level N fertilizer amounts of 7 crops (Maize, Wheat, Rice, Sugar cane, Rapeseed, Tea, Sugar beet) with IFASTAT data for 59 countries (Figure S1). Out of these 7 crops, Maize receives the maximum amount of N fertilizers, representing 20 % of global N fertilizer used on all the crops, followed by wheat with 18% and rice 16 % (IFA report, 2022). These three crops alone represent more than 50% of the total N fertilizer used globally. Our estimates show statistically significant correlation with IFASTAT data inventory with an R^2 of 0.94.

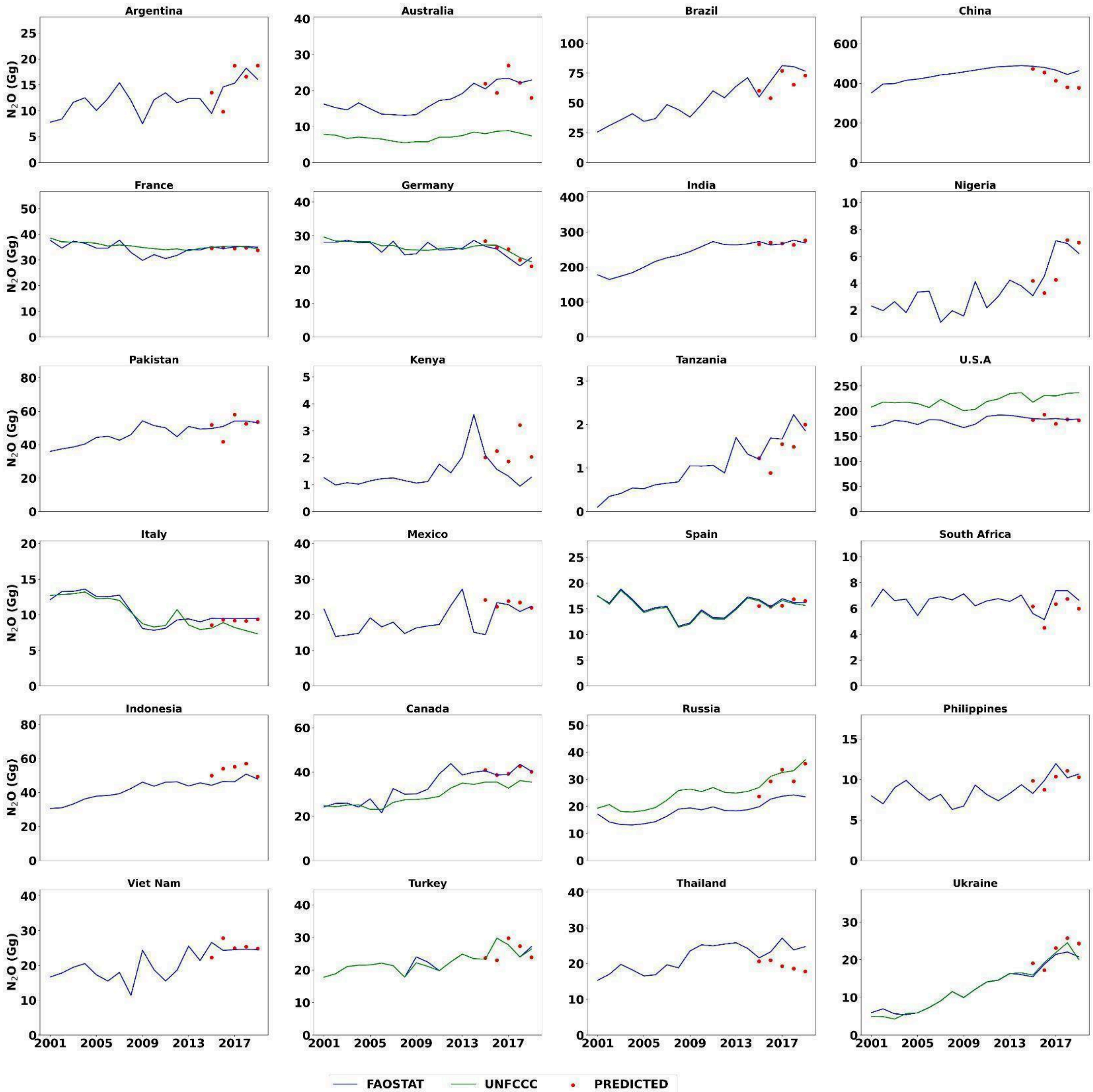

Figure 12 Time-series comparison of model-predicted N_2O emissions (red-dots) with FAO (blue-line) and UNFCCC inventory (green-line). Units- y-axis: Gigagram; x-axis: Year. Note: y-scales are different for each country. For the x-axis, FAO and UNFCCC data is plotted for the time period 2001-2019 and the model calculated emissions are for the time period 2016-2019.

3.2 Gridded emission estimate using IPCC tier 1 EF (~10 km)

In this section, we will explore the gridded N_2O emission map with an approximate resolution of 10 kilometers, created using the methodology outlined in section 2.2.5. This map is part of our broader effort to produce crop-specific N_2O emission maps for each year from 2015 to 2022 hosted on the Climate TRACE website, https://climatetrace.org/. Here, we focus on the cumulative emissions for all crops within each grid cell for the year 2021. An example of 2021 grid-level data is displayed in Figure 13. The map readily highlights emission hotspots, enabling us to pinpoint regions of intense N_2O output. Notably, the Gangetic Plain in India emerges as a significant hotspot, reflecting its status as one of the most agriculturally productive regions in the world. In the United States, areas such as the Corn Belt and the California Valley stand out as prominent emission centers. These regions are known for their intensive agricultural activities, which correlate with higher N_2O emissions. By identifying these hotspots, we can better target mitigation strategies and understand the regional contributions to global N2O emissions. Furthermore, to delineate the contribution of individual crops to the total emissions, we categorized the crops into four types: maize (depicted in red), rice (yellow), wheat (green), and other crops (grey), both at continental and global scales. Globally, maize, wheat, and rice are the predominant contributors, accounting for 80% of total emissions. On the continental scale, the patterns vary: in the Americas, maize is the primary contributor, responsible for 63% of emissions. In Asia, rice stands out as the leading emitter, although wheat (26%) and maize (22%) also have substantial shares. Europe's emissions are mostly dominated by wheat and a diverse array of other crops, together accounting for approximately 85% of the continent's total emissions. In contrast, Africa presents a more balanced emission profile, with wheat, maize, and other crops each having a roughly equivalent impact on the total N2O output. This nuanced understanding of crop-specific contributions is vital for formulating targeted emission reduction strategies across different regions. Identifying emission hotspots is crucial for effectively tackling N2O emissions from agriculture, as it allows for a focused approach to mitigation. By zeroing in on areas with the highest emissions, resources and strategies can be deployed more efficiently, ensuring the greatest possible impact on reducing greenhouse gases. This precise targeting not only makes economic sense but also facilitates the crafting of localized solutions that respect the nuances of individual agricultural systems, crop types, and environmental conditions.

3.3 Gridded emission estimate using heterogeneous EF (~4 km)

To further improve our emission estimate, we developed the spatio-temporal varying EF map (as discussed in section 2.2.6. To convert these EF maps into emission maps, we used the following step:

3.3.1 Rescaling ~10 km N fertilizer map to ~4 km spatial resolution

In our initial gridded dataset, we determined crop-specific nitrogen (N) fertilizer quantities at roughly 10 km resolution. For our crop-specific emission factor (EF) maps, developed using a novel method, we achieved a finer resolution of about 4 km. To synchronize the N fertilizer data with the EF maps, we started by calculating the N fertilizer rate (kg/ha) at the original resolution, which involved dividing the N fertilizer quantity by the grid area. We then resampled the N fertilizer rate maps to the finer 4 km resolution using a tool within Google Earth Engine (GEE). After obtaining these finer resolution maps of N fertilizer rates, we multiplied them by the grid area corresponding to the 4 km grids to reconstitute the N fertilizer amounts at this enhanced resolution. Finally, we multiplied these N fertilizer maps by the EF map to produce gridded, crop-specific emission maps. For comparison, we also applied the IPCC Tier 1 EF (1%) to the N fertilizer maps to generate alternate emission maps. Figure 14 illustrates the comparison of emission maps for crops such as maize, rice, and wheat using our new method versus the IPCC EF (1%) across the USA, India, and China.

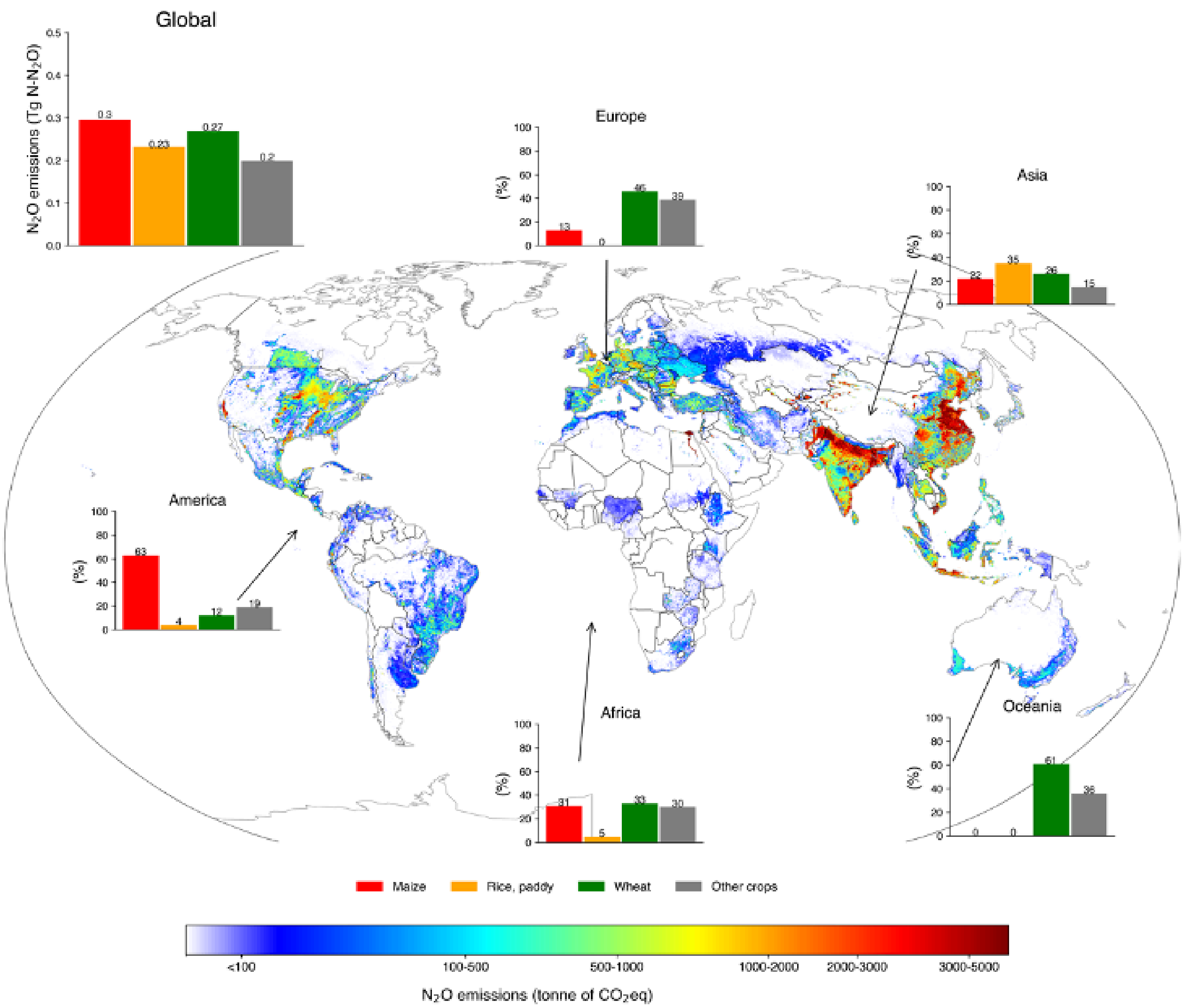

Figure 13 Total nitrogen fertilizer applied at grid scale 0.083° resolution (~10 km) for 2023. Units: Kg/ha grid area.

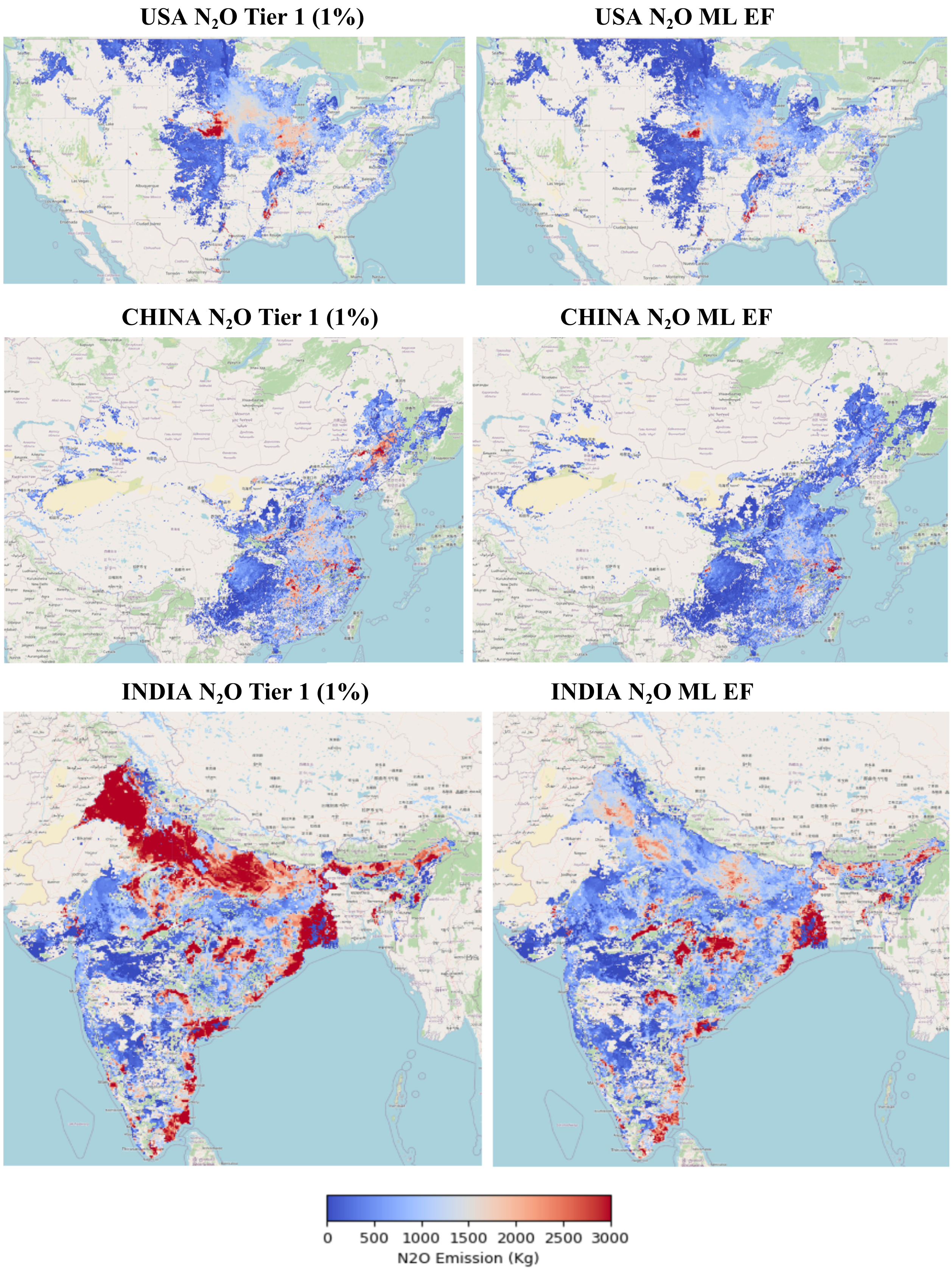

Figure 14 Comparison of N_2O emissions based on IPCC Tier 1 EF and our approach.

Table 7 Emission comparison between IPCC Tier 1 and the new method (italicized) used for USA, India, and China for 3 major crops: maize, rice and wheat in 2021.

Country	USA 2021			CHINA 2021			INDIA 2021		
Crop	Maize	Rice	Wheat	Maize	Rice	Wheat	Maize	Rice	Wheat
Emission IPCC (Gg)	95	0.97	9	96	65	48	5.6	110	78
Emission NEW EF (Gg)	*67*	*0.86*	*5*	*52*	*45*	*30*	*5.7*	*84*	*36*

4. Conclusion

In this study, we developed a novel data-based model to estimate direct N2O emissions from synthetic nitrogen fertilizer use on a country scale. This model assumes that crop nitrogen uptake is proportional to fertilizer application. Our estimates covered 103 countries, accounting for approximately 99% of global synthetic nitrogen fertilizer usage. From 2015 to 2020, total direct N_2O emissions were recorded as follows: 1,623, 1,610, 1,636, 1,585, 1,600, and 1,652 gigagrams (Gg) respectively. The seven largest contributors to these emissions were China, India, the USA, Brazil, Pakistan, Indonesia, and Canada, jointly responsible for over 60% of the total (>900 Gg annually). At the continental level, Asia, America, and Europe emerged as significant emitters, collectively accounting for 95.5% (1,542 Gg) of the global emissions averaged over the study period.

In the next phase, we refined our approach by estimating crop-specific emissions using gridded harvest areas and nitrogen fertilizer rates. This refined approach sheds more light on regional emission hotspots. For instance, India's Gangetic Plains and the USA's Corn Belt are identified as major contributors to their respective countries' emissions.

To further enhance our estimates, we integrated experimental emission factors (EF) from literature and global networks with remote-sensing data (including soil and climate factors) into machine learning algorithms. This allowed us to develop crop-specific, heterogeneous EFs that are more dynamic than the static 1% IPCC Tier 1 EF. Our method revealed significant differences in emission patterns when compared to those estimated using the IPCC's method. For example, the Gangetic Plains in India, which are identified as a major hotspot by the IPCC, were less prominent in our model. Additionally, our approach generally resulted in lower country-scale emissions estimates compared to those of the IPCC. Our goal is to enhance the precision of emission estimates using advanced tools like machine learning and remote sensing. We validated our estimates against various datasets to ensure reliability. Although our methodology differs significantly from that of the FAO or UNFCCC, it shows good correlation with other global emission inventories. This validation process builds more trust in current emissions inventories while identifying emerging inconsistencies. Our next step involves a dual-focus approach. Firstly, we plan to enhance our methodology by integrating additional features, such as a process-based model encompassing elements like nitrate, ammonium, and biomass. This inclusion will aid Machine Learning in better understanding the variability of emissions. Secondly, we will explore more experimental datasets in data-deficient regions such as Africa and South America. This ongoing improvement will progressively enhance our confidence in emission estimates.

Another area for potential improvement is reducing uncertainties related to crop yield and harvest area data. Currently, we rely on FAOSTAT, which is self-reported by countries. This reliance can be problematic, as countries might underreport or overreport their agricultural areas. For instance, Seto et al. (2000) highlighted significant discrepancies in China between satellite-derived and government-reported agricultural land area. The satellite-derived estimates for total agricultural land in ten counties were 115% greater than those reported in government yearbooks. This example illustrates the need for more reliable data sources in areas like sub-Saharan Africa, where national standards for agricultural data collection are often inadequate (Carletto et al., 2013).

5. Acknowledgements
This study was funded by Climate TRACE.

6. Supplementary materials

S1.1 N_2O emission maps for Maize at global scale using process-based model SALUS

In the latest update from last year, 2023, we integrated the process-based SALUS model (Basso et al., 2006), which simulates coupled plant and soil biogeochemistry processes to estimate N_2O emissions from maize cropping systems. The model considers three sources of emissions: synthetic nitrogen fertilizer, crop residue, and manure.

To run SALUS on a global scale, various datasets were utilized:

Table S1 Input data used for SALUS

Type	Source	Link
Weather	Prediction Of Worldwide Energy Resources (NASA/POWER)	datalink
Soil	SoilGrids	datalink
Synthetic N fertilizer	Section 2.2.5	
Manure N	Published study	datalink
Crop residue	Published study	datalink

Type	Source	Link
Tillage	Published study	datalink
Irrigation	AQUASTAT (FAOSTAT)	
Harvest Area	Spatial Production Allocation Model (SPAM)	datalink
Growing season	Published study	datalink

S1.2 Emission estimates

To estimate emission from each source synthetic N, manure, and crop residue from maize cropping systems,eight SALUS experiments were run with different management scenarios to separate total emission into different categories (Table S5).

In the first step, the method as discussed in the irrigation description of section 2.3.2.3 was used and we condensed the emission output from above experiments in Table S5. We used the following equations to obtain a combined emission estimate for each pair (Table S6). For example, for calculating emissions for Experiment type "*Baseline*", experimental ID (ExpID) pairs 1 and 5 were combined into a *Baseline* $N_2O_{combined}$ emissions. Similarly, this was done for the following ExpIDs-for {2,6} - $N_2O_{combined}$ (*No N fertilizer*); {3,7} - $N_2O_{combined}$ (*No manure*)' and{4,8} - $N_2O_{combined}$ (*No residue*). After these calculations, emission maps were created, representing one for each pair (Table S6; Figures S3 and S4).

In next step, to calculate the contribution of each sector (synthetic N, manure, and residue), we subtract the emission from the particular sector to the baseline:

$$N_2O_{synthetic\ N} = N_2O_{combined}\ (Baseline) - N_2O_{combined}\ (\text{No N fertilizer})$$

$$N_2O_{manure} = N_2O_{combined}\ (Baseline) - N_2O_{combined}\ (\text{No}\ manure)$$

$$N_2O_{residue} = N_2O_{combined}\ (Baseline) - N_2O_{combined}\ (\text{No}\ Residue)$$

Note, this dataset is currently in development and only emissions from maize systems have emissions estimates from 2023 to March 2024. This data is available on request by emailing: coalition@climatetrace.org.

S1.2 Data generated

In our latest update, we produced monthly emission data at a global scale for the maize cropping system, including emissions from synthetic nitrogen fertilizer, manure, and crop residue. This dataset features a spatial resolution of approximately 10 km, covering maize-growing grids worldwide. Temporally, the data spans from January 2023 to March 2024, providing detailed monthly estimates. Figures S3 and S4 present examples of the generated emission data, illustrating its coverage and variability.

Table S2 Total direct N_2O emissions due to synthetic nitrogen fertilizers for 5 continents for years 2015-2020 (all values rounded to one decimal place). Note, only countries modeled are included in the continent estimates.

	Total emission (Gg)					
Continent	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Asia	985.7	969.9	954.2	910.0	918.9	960.5
America	349.5	348.4	363.3	362.1	362.5	378.0
Europe	219.9	226.9	239.3	238.0	240.1	239.1
Africa	46.4	45.8	52.4	52.6	60.0	54.4
Oceania	21.9	19.4	27.0	22.2	18.0	19.7

Table S3 GAEZ crop classification

1. Wheat	10. Yams and other roots	19. Olives
2. Rice	11. Sugarbeet	20. Cotton
3. Maize	12. Sugarcane	21. Tobacco
4. Sorghum	13. Pulses	22. Banana
5. Millet	14. Soybean	23. Stimulants
6. Barley	15. Rapeseed	24. Vegetables
7. Other cereals	16. Sunflower	25 CropsNES
8. Potato and Sweet potato	17. Groundnut	26 Fodder crops
9. Cassava	18. Oil Palm fruit	

Table S4 Comparison for the percent change in emissions at country scale between the 5 years period from 2015 to 2019 for this study and other inventories.

Country	% Change in emissions from 2015 to 2019		
	Our study	**FAOSTAT**	**UNFCCC**
Argentina	38.61	69.35	Not available
Australia	-18.03	12.05	-7.22
Brazil	21.21	39.39	Not available
China	-20.21	-4.63	Not available
France	-2.25	-0.68	-1.55
Germany	-26.23	-12.36	-18.24
India	4.08	-1.53	Not available
Nigeria	68.15	101.98	Not available
Pakistan	3.30	6.90	Not available
Kenya	0.91	-38.94	Not available
Tanzania	63.24	54.7	Not available
U.S.A	-0.35	-0.51	8.64
Italy	9.54	-0.48	-9.85
Mexico	-9.08	55.15	Not available
Spain	6.23	-3.41	-5.38
South Africa	-2.87	18.30	Not available
Indonesia	-1.24	8.25	Not available
Canada	-1.83	-0.54	-0.03
Russian Federation	51.12	18.73	37.71
Philippines	4.55	29	Not available
Vietnam	11.65	-7.91	Not available
Turkey	0.91	16.35	13.18
Thailand	-13.75	14.55	Not available
Ukraine	27.51	34.27	25.54

Table S5 List of experiments - Synthetic N fertilizers, manure, and residue - performed with SALUS for maize cropping systems.

ExpID	Syn. N Fertilizer	Manure	Residue	Irrigation
1	Yes	Yes	Yes	Yes
2	No	Yes	Yes	Yes
3	Yes	No	Yes	Yes
4	Yes	Yes	No	Yes
5	Yes	Yes	Yes	No
6	No	Yes	Yes	No
7	Yes	No	Yes	No
8	Yes	Yes	No	No

Table S6 Combining experiment pair for irrigated and non-irrigated outputs maize cropping systems

Experiment type	ExpID Pair	Equation
Baseline	{1,5}	$N_2O_{combined}$ =(irrifac×N_2O_{irrig}($ExpID$ 1))+((1−irrifac)× $N_2O_{noirrig}$($ExpID$ 5)
No N fertilizer	{2,6}	$N_2O_{combined}$ =(irrifac×N_2O_{irrig}($ExpID$ 2))+((1−irrifac)× $N_2O_{noirrig}$($ExpID$ 6)
No Manure	{3,7}	$N_2O_{combined}$ =(irrifac×N_2O_{irrig}($ExpID$ 3))+((1−irrifac)× $N_2O_{noirrig}$($ExpID$ 7)
No Residue	{4,8}	$N_2O_{combined}$ =(irrifac×N_2O_{irrig}($ExpID$ 4))+((1−irrifac)× $N_2O_{noirrig}$($ExpID$ 8)

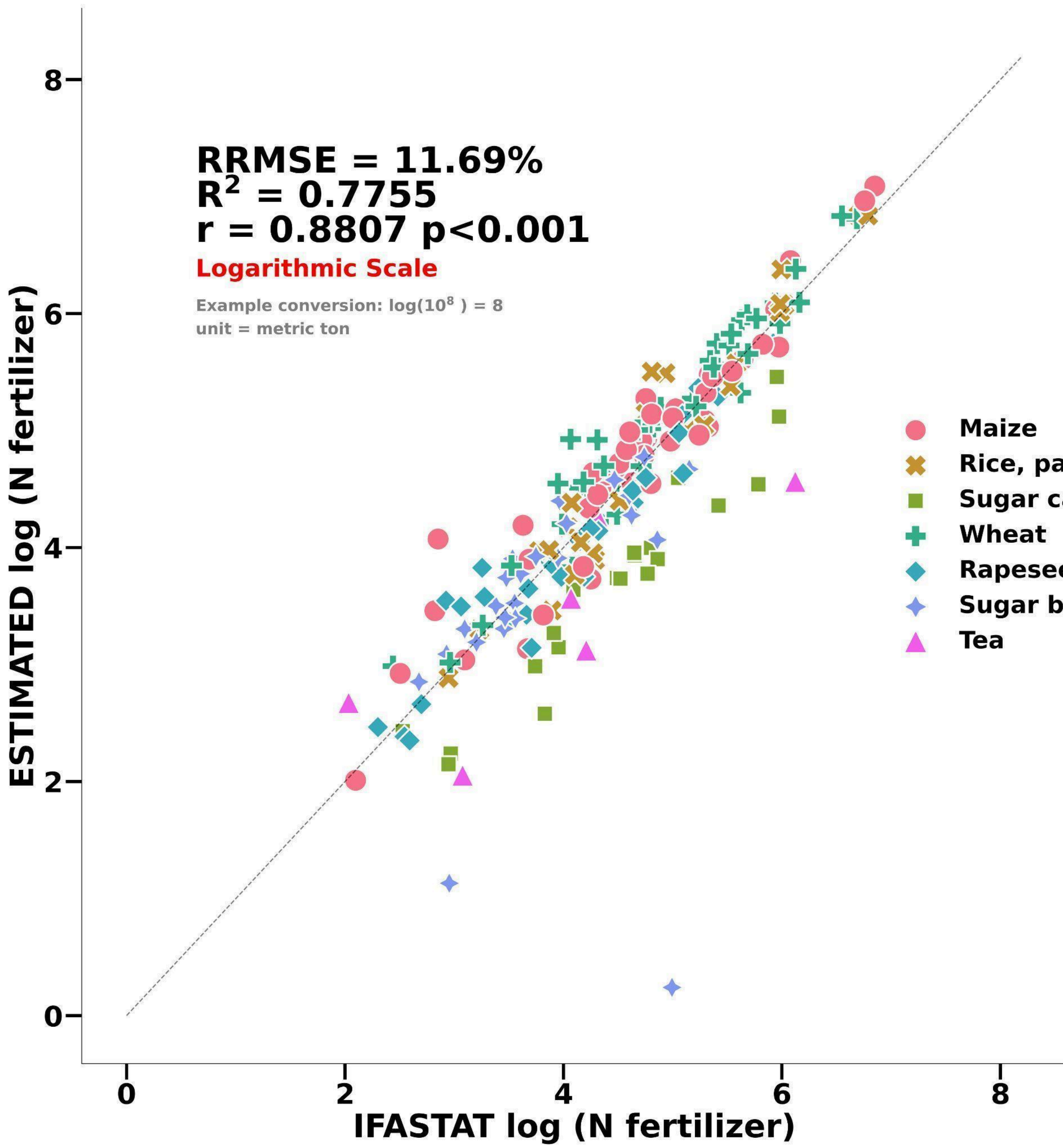

Figure S1 Comparison of crop-specific Nitrogen fertilizer at country scale between our study and IFASTAT for 2018. Units- y-axis and x-axis: kiloton (1000 metric ton).

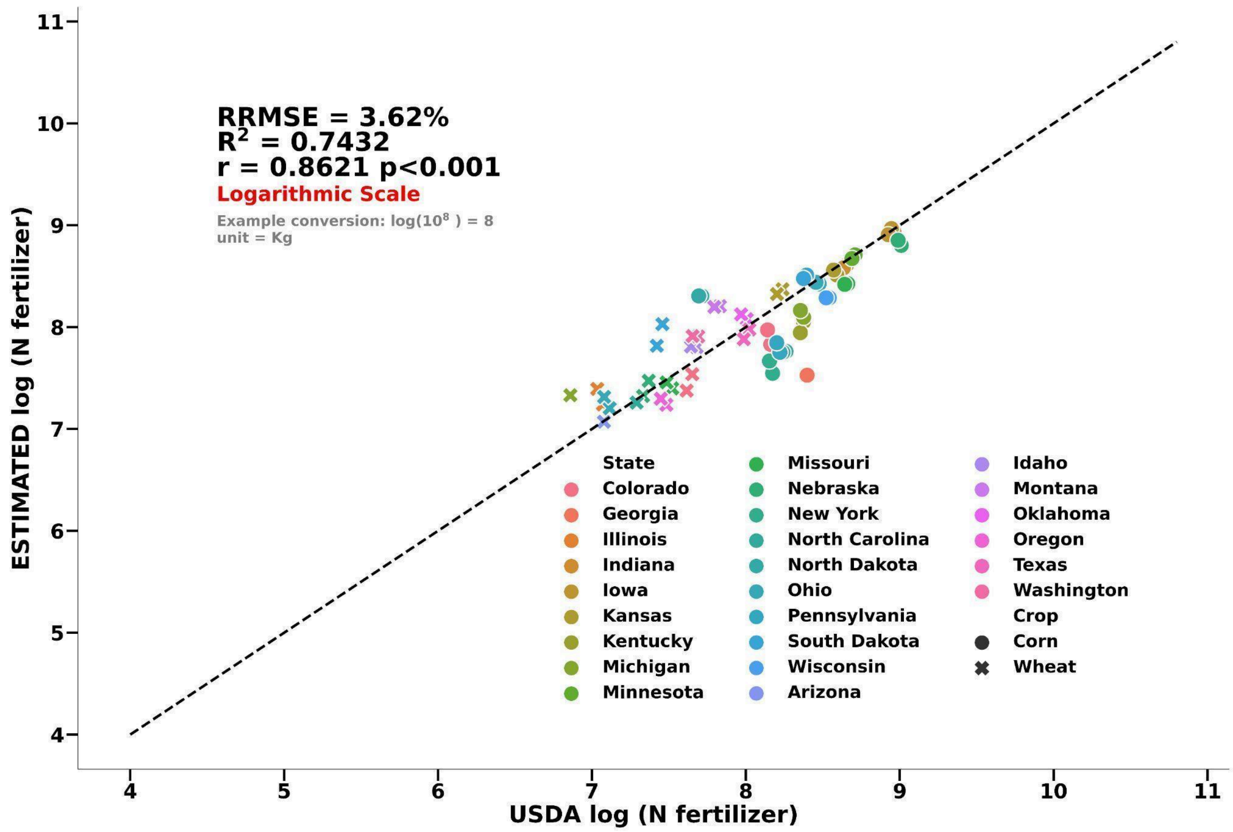

Figure S2 Comparison of crop-specific Nitrogen fertilizer at state scale between our study and USDA for 2015-2018.

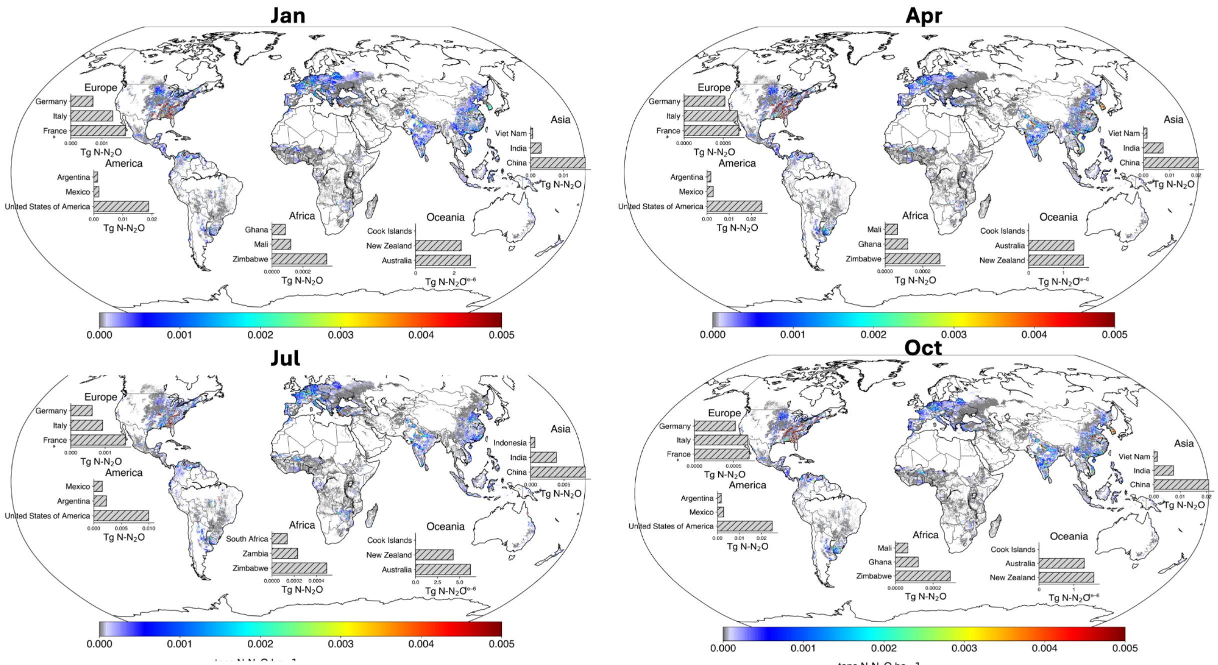

Figure S3 Monthly emission map for 2023 from synthetic N fertilizers for months January, April, July, and October for maize cropping systems**.**

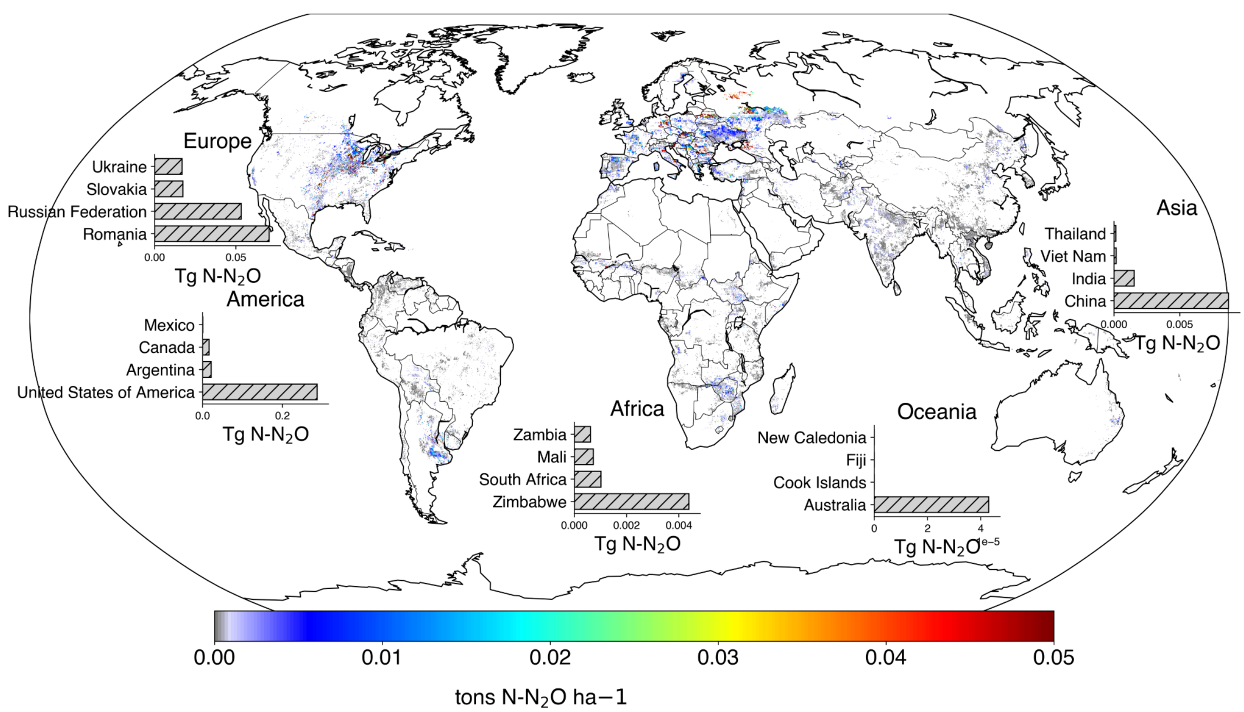

Figure S4 Annual emission map for 2023 from synthetic N fertilizers (top) and crop residue (bottom) for maize cropping systems**.**

Supplementary metadata section

Crop-specific direct N_2O emission are reported at two different spatial resolution (country and asset level) on the Climate TRACE website:

- Country scale crop-specific (include 140 crops in total) direct N_2O emission due to synthetic N fertilizer use in the croplands at 20 and 100 years GWPs
- Asset level crop-specific (140 crops combined in 26 types) direct N_2O emission due to synthetic N fertilizer use in the croplands at 20 and 100 years GWPs

Emissions estimates were reported for years 2015 to 2022, with 2015 backfilled with 2016 data. This sector does not include direct emissions produced from N fertilizer use within pasture. All data is freely available on the Climate TRACE website (https://climatetrace.org/). A detailed description of what is available is described in Table S1 and S2.

Table S4 General dataset information for country scale emissions.
file name: country-climate-trace_synthetic-fertilizer-application.csv.

General Description	Definition
Sector definition	*Crop-specific direct N_2O emission*
UNFCCC sector equivalent	*3.D.1.1 Inorganic N fertilizers*
Temporal Coverage	*2015 – 2023*
Temporal Resolution	*Annual (original); Monthly (on website, see Temporal Disaggregation of Emissions Data for the Climate TRACE Inventory)*
Data format(s)	*CSV and GeoTIFFs at ~10 km spatial resolution*
Coordinate Reference System	*None. ISO3 country code provided*
Number of countries available for download and percent of global emissions (as of 2022)	*106 total countries emission representing ~99% of this sector's emissions*
Ownership	*Country*
What emission factors were used?	*IPCC tier 1*
total_CO2e_100yrGWP and total_CO2e_20yrGWP conversions	*Climate TRACE uses IPCC AR6 CO_2e GWPs. CO_2e conversion guidelines are here:https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_FullReport_small.pdf*

Table S5 Source level metadata description confidence and uncertainty.

Data attribute	Confidence Definition	Uncertainty Definition
type	• *Medium*: if estimate is modeled and not validated	

Data attribute	Confidence Definition	Uncertainty Definition
	● *High*: if estimate type is modeled and validated	Not used; N/A
capacity_description	● *Medium*: if estimate type was modeled based on published data	Given as an interval with an lower and upper bound of value
activity_description	● *Medium*: if estimate is modeled and not validated ● *High*: if estimate type is modeled and validated for crop-specific N fertilizer amount	Given as an interval with an lower and upper bound of value
CO2_emissions_factor	Not used; N/A	Not used; N/A
CH4_emissions_factor	Not used; N/A	Not used; N/A
N2O_emissions_factor	*Medium*: based on IPCC emissions factors	IPCC uncertainty estimates, expressed as a percentage above or below the mean estimate (i.e. +/-XX%), or as an interval with an upper and lower bound of values.
other_gas_emissions_factor	Not used; N/A	Not used; N/A
CO2_emissions	Not used; N/A	Not used; N/A
CH4_emissions	Not used; N/A	Not used; N/A
N2O_emissions	● *Medium*: if estimate is modeled and not validated ● *High*: if estimate type is modeled and validated total amount with UNFCCC	Given as an interval with an lower and upper bound of value
other_gas_emissions	Not used; N/A	Not used; N/A
total_CO2e_100yrGWP	● *Medium*: if estimate is modeled and not validated ● *High*: if estimate type is modeled and validated total amount with UNFCCC	Given as an interval with an lower and upper bound of value
total_CO2e_20yrGWP	● *Medium*: if estimate is modeled and not validated ● *High*: if estimate type is modeled and validated total amount with UNFCCC	Given as an interval with an lower and upper bound of value

Permissions and Use: All Climate TRACE data is freely available under the Creative Commons Attribution 4.0 International Public License, unless otherwise noted below.

Data citation format: Sharma, P. and Basso, B. (2024). *Agriculture sector- Estimation of Direct Nitrous Oxide (N2O) Emissions from Synthetic Fertilizers*. Department of Earth and

Environmental Sciences, Michigan State University, USA, Climate TRACE Emissions Inventory. https://climatetrace.org [Accessed date]

Geographic boundaries and names (iso3_country data attribute): The depiction and use of boundaries, geographic names and related data shown on maps and included in lists, tables, documents, and databases on Climate TRACE are generated from the Global Administrative Areas (GADM) project (Version 4.1 released on 16 July 2022) along with their corresponding ISO3 codes, and with the following adaptations:

- HKG (China, Hong Kong Special Administrative Region) and MAC (China, Macao Special Administrative Region) are reported at GADM level 0 (country/national);
- Kosovo has been assigned the ISO3 code 'XKX';
- XCA (Caspian Sea) has been removed from GADM level 0 and the area assigned to countries based on the extent of their territorial waters;
- XAD (Akrotiri and Dhekelia), XCL (Clipperton Island), XPI (Paracel Islands) and XSP (Spratly Islands) are not included in the Climate TRACE dataset;
- ZNC name changed to 'Turkish Republic of Northern Cyprus' at GADM level 0;
- The borders between India, Pakistan and China have been assigned to these countries based on GADM codes Z01 to Z09.

The above usage is not warranted to be error free and does not imply the expression of any opinion whatsoever on the part of Climate TRACE Coalition and its partners concerning the legal status of any country, area or territory or of its authorities, or concerning the delimitation of its borders.

Disclaimer: The emissions provided for this sector are our current best estimates of emissions, and we are committed to continually increasing the accuracy of the models on all levels. Please review our terms of use and the sector-specific methodology documentation before using the data. If you identify an error or would like to participate in our data validation process, please contact us.

7. References

1. FAO (Food and Agriculture Organization of the United Nations). FAO statistical databases. (2019) Available at: http://faostat.fao.org.
2. IFA (International Fertilizer Association). IFA statistical databases.(2022) Available at https://www.ifastat.org.
3. Carletto C., Jolliffe D. and Banerjee R. (2013). The emperor has no data! agricultural statistics in Sub-Saharan Africa World Bank Working Paper. http://mortenjerven.com/wp-content/uploads/2013/04/Panel-3-Carletto.pdf
4. Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J. G. J., Peters, J. A. H. W., van Aardenne, J. A., Monni, S., Doering, U., Petrescu, A. M. R., Solazzo, E., & Oreggioni, G. D. (2019). EDGAR

v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970–2012. Earth System Science Data, 11(3), 959–1002. https://doi.org/10.5194/essd-11-959-2019

5. Lassaletta, L., Billen, G., Grizzetti, B., Anglade, J., & Garnier, J. (2014). 50 year trends in nitrogen use efficiency of world cropping systems: the relationship between yield and nitrogen input to cropland. Environmental Research Letters, 9(10), 105011. https://doi.org/10.1088/1748-9326/9/10/105011
6. Liu, J., You, L., Amini, M., Obersteiner, M., Herrero, M., Zehnder, A. J. B., & Yang, H. (2010). A high-resolution assessment on global nitrogen flows in cropland. Proceedings of the National Academy of Sciences, 107(17), 8035–8040. https://doi.org/10.1073/pnas.0913658107
7. Northrup, D. L., Basso, B., Wang, M. Q., Morgan, C. L. S., & Benfey, P. N. (2021). Novel technologies for emission reduction complement conservation agriculture to achieve negative emissions from row-crop production. Proceedings of the National Academy of Sciences, 118(28), e2022666118. https://doi.org/10.1073/pnas.2022666118
8. Seto, K. C., Kaufmann, R. K., & Woodcock, C. E. (2000). Landsat reveals China's farmland reserves, but they're vanishing fast. *Nature*, *406*(6792), 121–121. https://doi.org/10.1038/35018267
9. Syakila, A., & Kroeze, C. (2011). The global nitrous oxide budget revisited. Greenhouse Gas Measurement and Management, 1(1), 17–26. h7ttps://doi.org/10.3763/ghgmm.2010.000
10. Tian, H., Xu, R., Canadell, J. G., Thompson, R. L., Winiwarter, W., Suntharalingam, P., Davidson, E. A., Ciais, P., Jackson, R. B., Janssens-Maenhout, G., Prather, M. J., Regnier, P., Pan, N., Pan, S., Peters, G. P., Shi, H., Tubiello, F. N., Zaehle, S., Zhou, F., . . . Yao, Y. (2020). A comprehensive quantification of global nitrous oxide sources and sinks. Nature, 586(7828), 248–256. https://doi.org/10.1038/s41586-020-2780-0
11. Thompson, R. L., Lassaletta, L., Patra, P. K., Wilson, C., Wells, K. C., Gressent, A., Koffi, E. N., Chipperfield, M. P., Winiwarter, W., Davidson, E. A., Tian, H., & Canadell, J. G. (2019). Acceleration of global N_2O emissions seen from two decades of atmospheric inversion. Nature Climate Change, 9(12), 993–998. https://doi.org/10.1038/s41558-019-0613-7
12. Tubiello, F. N., Rosenzweig, C., Conchedda, G., Karl, K., Gütschow, J., Xueyao, P., Obli-Laryea, G., Wanner, N., Qiu, S. Y., Barros, J. D., Flammini, A., Mencos-Contreras, E., Souza, L., Quadrelli, R., Heiðarsdóttir, H. H., Benoit, P., Hayek, M., & Sandalow, D. (2021). Greenhouse gas emissions from food systems: building the evidence base. Environmental Research Letters, 16(6), 065007. https://doi.org/10.1088/1748-9326/ac018e
13. UNESCAP, UNECA, UNECE, and UN ESCWA, U., 2020. World economic situation and prospects. Available at: https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/WESP2020_FullReport.pdf
14. Vitousek, P. M., Naylor, R., Crews, T., David, M. B., Drinkwater, L. E., Holland, E., Johnes, P. J., Katzenberger, J., Martinelli, L. A., Matson, P. A., Nziguheba, G., Ojima, D., Palm, C. A., Robertson, G. P., Sanchez, P. A., Townsend, A. R., & Zhang, F. S. (2009). Nutrient

Imbalances in Agricultural Development. Science, 324(5934), 1519–1520. https://doi.org/10.1126/science.1170261

15. Winiwarter, W., Höglund-Isaksson, L., Klimont, Z., Schöpp, W., & Amann, M. (2018). Technical opportunities to reduce global anthropogenic emissions of nitrous oxide. Environmental Research Letters, 13(1), 014011. https://doi.org/10.1088/1748-9326/aa9ec9