



HAL
open science

Sources of uncertainty in simulating crop N₂O emissions under contrasting environmental conditions

Sibylle Dueri, Joël Léonard, Florent Chlebowski, Pablo Rosso, Michael Berg-Mohnicke, Claas Nendel, Fiona Ehrhardt, Pierre Martre

► To cite this version:

Sibylle Dueri, Joël Léonard, Florent Chlebowski, Pablo Rosso, Michael Berg-Mohnicke, et al.. Sources of uncertainty in simulating crop N₂O emissions under contrasting environmental conditions. *Agricultural and Forest Meteorology*, 2023, 340, pp.109619. 10.1016/j.agrformet.2023.109619. hal-04264455

HAL Id: hal-04264455

<https://hal.inrae.fr/hal-04264455v1>

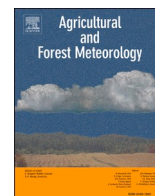
Submitted on 30 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Sources of uncertainty in simulating crop N₂O emissions under contrasting environmental conditions

Sibylle Dueri^{a,*}, Joël Léonard^b, Florent Chlebowski^b, Pablo Rosso^c, Michael Berg-Mohnicke^c, Claas Nendel^{c,d}, Fiona Ehrhardt^e, Pierre Martre^{a,**}

^a LEPSE, Univ Montpellier, INRAE, Institut Agro Montpellier SupAgro, Montpellier, France

^b BioEcoAgro Joint Research Unit, INRAE, Université de Liège, Université de Lille, Université de Picardie Jules Verne, Barenton-Bugny, France

^c Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany

^d University of Potsdam, Institute of Biochemistry and Biology, Potsdam, Germany

^e RITTIMO Agroenvironment, Colmar, France

ARTICLE INFO

Keywords:

Agroecosystem model
Uncertainty
Denitrification
Model intercomparison
Nitrification
N₂O emission

ABSTRACT

Nitrogen fertilization is a key agronomic lever for high crop productivity, but also an important source of N₂O emission, a potent greenhouse gas. Process-based agroecosystem simulation models are popular tools for managing the timing and amount of fertilization, and help reduce N₂O emissions. However, accurate simulation of N₂O emissions at field scale is still a challenge due to the spatial and temporal variability of the soil conditions. In this study, we investigated the sources of structural uncertainty in predicting N₂O emissions under a wide range of pedo-climatic conditions using a representative field data set. We implemented the same nitrification/denitrification/N₂O emission formalism in three different agroecosystem models and analyzed how the inter-model variability of variables involved in nitrification and denitrification processes, affected the simulated N₂O emissions. We characterized the dispersion of the key variables (water-filled pore space, NO₃⁻ and NH₄⁺ concentration, and soil temperature) between models and we evaluated the effect of variable uncertainty on N₂O emissions uncertainty using a sensitivity analysis. We also analyzed model errors over a wide range of soil-climate conditions to identify the most challenging conditions for simulation, which require further model improvement. Our results highlighted that the simulation of the timing and amplitude of the NO₃⁻ and NH₄⁺ peaks was highly variable between agroecosystem models, with an important impact on N₂O emission. These peaks occurred mainly after fertilization or incorporation of crop residues, and the different representations of fertilization and mineralization between the models had a major effect on the simulation of N₂O emissions. Our analysis also emphasized that wet acidic soils with high denitrification potential are more challenging for models to simulate.

1. Introduction

N₂O is a long-lived greenhouse gas (GHG) that accumulates in the atmosphere and that depletes stratospheric ozone (Ravishankara et al., 2009). Agriculture accounts for 52% of the anthropogenic emissions of N₂O, which are dominated by nitrogen fertilizer application to croplands (Tian et al., 2020). Given the strong warming potential of N₂O and the important contribution of agricultural soils in N₂O emissions, there is an urgent need to develop strategies that improve agricultural practices and reduce N₂O emission (Signor and Cerri, 2013).

From 2000 to 2018, N₂O emission from synthetic fertilizers and crop residues incorporation have increased by 35% due to growing worldwide intensification of crop production and increase in chemical fertilizers inputs (FAOSTAT, 2018). Estimation from global atmospheric inversion show an acceleration of N₂O emission during the last decade, especially in East Asia and South America (Thompson et al., 2019). The increase in N₂O emissions is expected to continue in the coming decades due to population growth and rising demand for agricultural production (Alexandratos and Bruinsma, 2012).

Two biogeochemical processes control N₂O emissions: nitrification

* Corresponding author at: Agroscope, Tänikon, Switzerland.

** Corresponding author at: LEPSE, INRAE, Montpellier France.

E-mail addresses: sibylle.dueri@admin.agroscope.ch (S. Dueri), pierre.martre@inrae.fr (P. Martre).

¹ Current address: Agroscope, Research Group Economic Modeling and Policy Analysis, 8356 Ettenhausen, Switzerland

and denitrification. Nitrification is the oxidation of NH_4^+ to NO_3^- , while denitrification is the reduction of NO_3^- to molecular nitrogen (N_2). Both processes are facilitated by soil micro-organisms and are affected by soil conditions, such as temperature, moisture, oxygen content, pH, and amount of available organic carbon and nitrogen (Butterbach-Bahl et al., 2013). A variable fraction of the nitrification and denitrification rates leaves the soil as N_2O emissions. To optimize nitrogen fertilizer application and reduce N_2O emission, we need to understand the interactions between soil conditions and N_2O production and how the biogeochemical processes are impacted by pedo-climatic conditions (Signor and Cerri, 2013).

Process-based agroecosystem models simulate the biogeochemical processes that affect the N fluxes in the soil (nitrification, denitrification, and mineralization), the N uptake by the crop and the loss of N to the atmosphere and to the groundwater under the influence of environmental factors and farmer's management (fertilization, tillage, irrigation, and organic matter amendments; Nendel et al., 2014). These models have been widely used to simulate N dynamics in croplands and to assess environmental impacts such as leaching, volatilization, and denitrification (Cannavo et al., 2008). They are valuable tools for improving the timing and amount of fertilization and help reduce unwanted N emission. However, accurate simulation of N_2O emissions at field scale is still challenging due to the spatial and temporal variability of the soil conditions, which affect the microbial processes underlying N_2O production (Chen et al., 2008; Fuchs et al., 2020; Berardi et al., 2020). The simulation of N_2O emissions can be very different from one model to another, especially the timing and magnitude of peak emissions (Fuchs et al., 2020; Gaillard et al., 2018). The limited availability of daily driving variable data is a major problem for model evaluation, as it is difficult to determine whether the error depends on a misrepresentation of the input data or on an error in the formalisms (Del Grosso et al., 2020).

During the last 20 years, there has been an increasing effort towards modularity of agricultural and ecological models, to improve model comparison and reusability of modules (Jones et al., 2001; Reynolds and Acock, 1997; Rizzoli et al., 2008). Agroecosystem models have been integrated into modeling platforms, which allow different model components to be assembled. Efforts have also been made to facilitate model component re-use and exchange between modeling platforms despite the different platform formalisms (Midingoyi et al., 2021) with the idea that the increasing accessibility of reusable modules, allows for testing of alternative modeling hypotheses and ultimately for model improvement (Muller and Martre, 2019; Berg-Mohnicke and Nendel, 2022). Each model component represents a set of bio-physical processes through mathematical formalisms and is connected to the other components through common variables that are exchanged at the interface of model components. The accuracy of the output of a model component does not only depend on the appropriateness of the formalisms used to represent the biophysical processes, but also on the accuracy of other connected modules. It is important to understand how and why the output of a model component changes when that same component is connected to different agroecosystem models, but to our knowledge, no study has yet investigated this question.

In this study, the N_2O emission module originally implemented in STICS (Brisson et al., 2010) was integrated in two different agroecosystem models: MONICA (Nendel et al., 2011) and *SiriusQuality* (Martre et al., 2006). The three agroecosystem models were used to simulate 31 year/location/treatment/soil combinations, representing pedo-climatic conditions of contrasting locations across the world. The objective was to assess the variability in the predictions of N_2O emission between agroecosystem models and identify the sources of these differences. In a first step, we characterized the uncertainty of the input variables (water-filled pore space [WFPS], NH_4^+ and NO_3^- concentration, soil temperature) between models. Then we used a sensitivity analysis to evaluate how this variability affected the simulated N_2O emission.

Finally, we compared the accuracy of the models with respect to N_2O emission under different pedo-climatic conditions, and identified conditions that may be more challenging to simulate, which will require further model improvement.

2. Material and methods

2.1. Experimental dataset and simulation

The experimental data set we used in this study is composed of six previously published arable crop field experiments that provided high-quality data on N_2O emission. The experimental sites are located in Ottawa, Canada (Jégo et al., 2012; Sansoulet et al., 2014), Grignon, France (Loubet et al., 2011), New Delhi, India (Bhatia et al., 2012), Kingaroy, Australia (De et al., 2014), Santa Maria, Brazil (Aita et al., 2015, 2019), and Estrées-Mons, France (Coudrain et al., 2016; Domeignoz-Horta et al., 2015; Domeignoz-Horta et al., 2018) and cover a wide range of soil, weather and crop management conditions (Table 1). Data from the first five locations were previously used in a multi-model intercomparison analysis (Ehrhardt et al., 2018).

The data set contains detailed information on crop management, weather conditions, soil properties, daily N_2O emission, and measurements over the growing season of soil NH_4^+ , NO_3^- and water content. It also contains in-season measurements of leaf area index, total above ground biomass and nitrogen and final grain yield and nitrogen. For Ottawa, Grignon, New Delhi, Kingaroy and Santa Maria, the data set contains one treatment per year. For Estrées-Mons (ACBB long term experiment), data were available for five treatments in 2013 and 2014 and 6 treatments in 2017 and 2019. Each of these treatments is characterized by a combination of nitrogen fertilization, crop residue management, tillage depth, and cover crop (Domeignoz-Horta et al., 2018). In Table 1 we have grouped the treatments according to nitrogen supply into low (L), intermediate (I) and high (H) N input treatment.

2.2. Nitrification, denitrification and N_2O emission module

The N_2O module computes the nitrification and denitrification fluxes, and the associated emission of N_2O . Nitrification and denitrification are coupled through NO_3^- concentration, which is produced by nitrification and used as substrate for denitrification. The original equations of the module were presented in Bessou et al. (2010) and later improved and used in Peyrard et al. (2017) and Plaza-Bonilla et al. (2017). A detailed description is available in Beaudoin et al. (2023).

Nitrification and N_2O emission - The nitrification rate is proportional to the NH_4^+ concentration and depends on pH, temperature, and soil water content (Fig. 1). The nitrification rate decreases with decreasing pH: it is optimum at $\text{pH} > 7.2$ and decreases to null at $\text{pH} = 4$. The effect of temperature is described by a Gaussian function, with an optimum at 32.5°C (Benoit et al., 2015). The rate of nitrification increases with soil water content up to field capacity. Beyond that, soil water limits soil aeration and decreases the nitrification rate (Kahlil et al., 2004).

The N_2O emission associated with nitrification is favored by anaerobic conditions, which depend on the soil water content. In the module, the fraction of nitrification that goes to N_2O emission tends to 0.16% at low soil water content and increases sharply for a WFPS of 0.8 to reach a maximum value of 2.52% when the soil is water saturated (Kahlil et al., 2004).

Denitrification and N_2O emission - The potential denitrification rate is either calibrated or can be estimated from the soil organic carbon concentration. The latter option was used in this study. Denitrification rate is obtained by multiplying the potential rate with weighting factors that account for the effect of NO_3^- concentration, soil temperature, and soil water content (Fig. 1). The effect of increasing NO_3^- concentration is to increase the denitrification rate following a Michaelis-Menten kinetics with a half saturation constant equal to $148\text{ mg NO}_3^- \cdot \text{N L}^{-1}$. The

Table 1
Experiment name, location, and meteorological conditions during the growing season, crop management, and soil characteristics in the 0–30 cm soil layer for the field experiments simulated in this study (SB, spring barley; WW, winter wheat; SW, spring wheat).

| Treatment code | Location, country | Latitude [°] | Longitude [°] | Weather conditions | | | Crop management | | | Soil characteristic | | | | |
|----------------|-------------------|--------------|---------------|--------------------------|-------------------------------|-------------------------------------------------|---------------------|--------------------|-------------|---------------------------------------------------------------------|-----------------------------------------------|-----------------|-----------------------------|------------------------------------|
| | | | | Average temperature [°C] | Cumulative precipitation [mm] | Cumulative solar radiation [MJ/m ²] | Number of treatment | Cultivar (species) | Sowing date | Total N fertilizer [kg N ha ⁻¹] (number of application) | Total irrigation [mm] (number of application) | pH in water [-] | Water filled pore space [-] | Organic C [g 100 g ⁻¹] |
| CA.07 | Ottawa, | 45.41 | -75.69 | 18.5 | 344 | 2351 | 1 | ACBrio | 2007-05-19 | 0 (0) | 0 (0) | 6.6 | 0.49 | 3.1 |
| CA.11 | Canada | | | 19.4 | 300 | 2284 | | (SW) | 2011-05-10 | 78 (1) | | | 0.41 | |
| GR.08 | Grignon, | 48.85 | 1.91 | 8.8 | 478 | 3117 | 1 | Premio | 2008-10-17 | 83.5 (2) | 0 (0) | 7.75 | 0.61 | 1.5 |
| GR.11 | France | | | 9.6 | 518 | 3110 | | (WW) | 2011-10-18 | 96 (2) | | | 0.74 | |
| IN.06 | New Delhi, | 28.64 | 77.21 | 17.1 | 112 | 1867 | 1 | PBW343 | 2006-11-29 | 120 (3) | 250 (5) | 8 | NA | 0.35 |
| IN.07 | Inde | | | 17.1 | 33 | 2231 | | (SW) | 2007-12-01 | | | | | |
| IN.08 | | | | 18.6 | 16 | 2829 | | | 2008-11-25 | | | | | |
| AU.11 | Kingaroy, | -26.53 | 151.83 | 15.7 | 176 | 2829 | 1 | Hartog | 2011-07-06 | 80 (2) | 136 (4) | 5.6 | 0.51 | 1.3 |
| | Australia | | | | | | | (WW) | | | | | | |
| BR.13 | Santa Maria, | -29.68 | -53.8 | 16.2 | 571 | 2237 | 1 | Quartzo | 2013-06-08 | 0 (0) | 0 (0) | 5.4 | 0.89 | 1.9 |
| | Brazil | | | | | | | (WW) | | | | | | |
| EM.13.H | Estrées- | 49.88 | 3 | 13.5 | 251 | 2290 | 5 | Sebastian | 2013-03-26 | 100 (1) | 0 (0) | 8.2 | 0.68 | 1.05 |
| EM.13.L | Mons, France | | | | | | | (SB) | | 40 (1) | | | | |
| EM.15.H | | | | 9.5 | 416 | 3040 | 5 | Cellule | 2014-10-30 | 170 (3) | | | 0.79 | |
| EM.15.L | | | | | | | | (WW) | | 60 (1) | | | | |
| EM.18.H | | | | 9.8 | 534 | 2898 | 6 | Absalon | 2017-10-18 | 200 (4) | | | 0.61 | |
| EM.18.L | | | | | | | | (WW) | | 80 (2) | | | | |
| EM.18.I | | | | | | | | | | 120 (2) | | | | |
| EM.19.H | | | | 14.7 | 147 | 2652 | 6 | RGT | 2019-03-25 | 160 (2) | | | 0.43 | |
| EM.19.L | | | | | | | | Planet | | 55 (1) | | | | |
| | | | | | | | | (SB) | | | | | | |

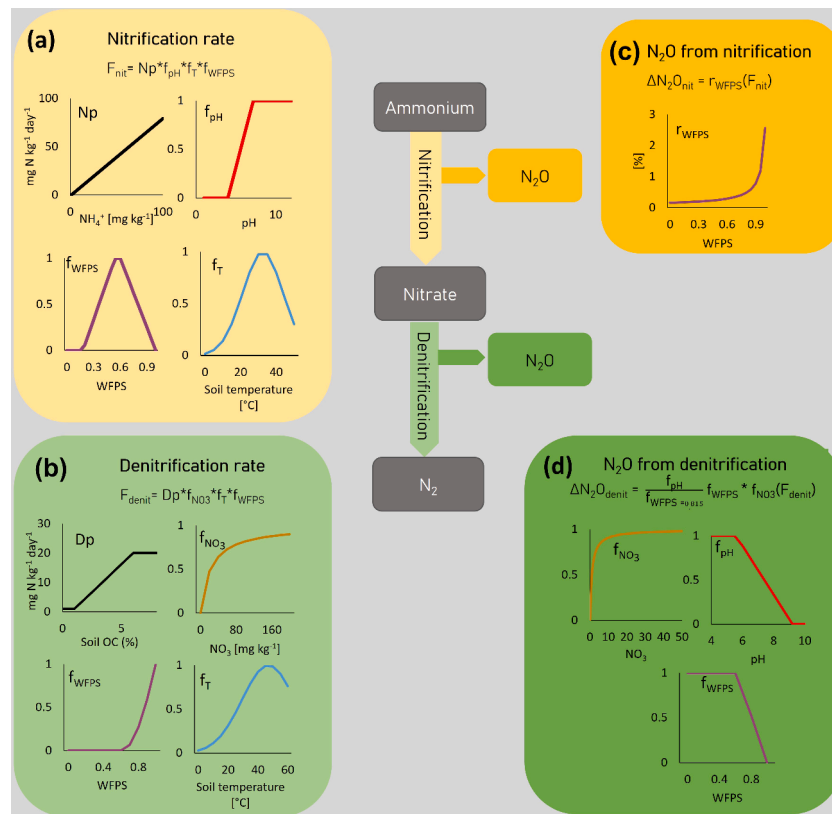


Fig. 1. Representation of the model of nitrification rate (a), denitrification rate (b) N_2O emission from nitrification (c) and N_2O emission from denitrification (d) with mathematical equations and graphical representation of the effect of NH_4^+ nitrate NO_3^- , pH, soil temperature, WFPS, and soil organic carbon (Soil OC). The colors of the boxes correspond to the fluxes identified in the conceptual diagram at the center.

weighting factor associated to soil temperature increases following a Gaussian function, with an optimum at 47 °C (Benoit et al., 2015). Denitrification rate is null when WFPS is smaller than 0.62, while it increases exponentially to reach a value of 1 when the soil is water saturated.

The N_2O emission associated with denitrification depends on pH and soil water content. Acid pH strongly inhibits N_2O reduction to N_2 (Rochester, 2003). In the N_2O module, the fraction of denitrification going to N_2O emission decreases linearly between pH 5.6 and 9.2. At high soil water contents, which result in more anaerobic conditions, the reduction of N_2O to N_2 is favored (Vieten et al., 2008). In the N_2O module, this effect is represented by a linear decrease of the fraction of denitrification going to N_2O emission between WFPS 0.62 and saturation.

2.3. Coupling of the N_2O module with the MONICA, SiriusQuality and STICS agroecosystem models

The agroecosystem models considered in this study (MONICA, SiriusQuality, and STICS) are process-based simulation models that represent the most important processes affecting the soil-plant system and that simulate the crop and soil variables on a daily time step. The models estimate simultaneously agricultural and environmental variables (e.g. soil water, crop biomass, crop nitrogen, soil nitrogen) by taking into account the impact of weather, soil, and management practice

2.3.1. MONICA

The MONICA model (version 3.3.1; Tendel et al., 2011) simulates the soil water, N, and C fluxes and temperature for each 0.1-m thick soil layers as defined to a maximum soil depth of 2 m (Table 2). Soil carbon dynamics is described by three conceptual pools (soil organic matter, microbial biomass and freshly added organic matter), each with a rapid

Table 2

Soil module characteristics (fluxes, pools, processes and management options) of the MONICA, SiriusQuality and STICS agroecosystem models.

| | Agroecosystem models | | |
|---------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|
| | MONICA | SiriusQuality | STICS |
| Soil fluxes | water, N, C | water, N | water, N, C |
| Vertical grid of the soil model | 0.1 m | 0.05 m (2 soil layer for soil temperature) | 0.01 m |
| C and N pools | - organic matter - microbial biomass - added organic matter | - organic N - added organic N | - organic residues - microbial biomass - humus - mineral pools |
| N processes | - mineralization - residue decomposition - volatilization - immobilization - leaching - nitrification - denitrification | - mineralization - residue decomposition - leaching - nitrification - denitrification | - mineralization - residue decomposition - volatilization - immobilization - leaching - nitrification - denitrification |
| Management options | - N fertilization - irrigation - residue incorporation - ploughing | - N fertilization - irrigation - residue incorporation | - N fertilization - irrigation - residue incorporation - soil tillage |

and a slow turn-over rate as borrowed from the DAISY model (Abrahamson and Hansen, 2000), and has been thoroughly tested by Aiteev et al. (under review). Besides mineralization, MONICA explicitly represents volatilization (NH_3 emissions, Søgaard et al., 2002) from manure, slurry and urea application, and N immobilization in the soil as a result of the microbial biomass dynamics. Different management

options are included in the model, including residue incorporation, ploughing and automated fertilization and irrigation. A generic crop module simulates the growth and development of a range of agricultural crops, including barley (Rötter et al., 2012), wheat (Asseng et al., 2013), maize (Bassu et al., 2014) soybean (Nendel et al., 2023) and potato (Fleisher et al., 2017), in rotations (Kollas et al., 2015).

2.3.2. *SiriusQuality*

The *SiriusQuality* model (version 3.0; Martre et al., 2006) simulates the soil N and water fluxes, but does not explicitly represent C fluxes. Each soil layer has a thickness of 5 cm and the model computes the N and water content by layer until the maximum rooting depth, as defined by the model user. Soil temperature is calculated for the near-surface layer and deep soil (Jamieson et al., 1995). The organic matter turnover is represented by a soil mineralization rate that depends on the physical and chemical conditions of the soil, as described in Clivot et al. (2017). If organic residues are added to the soil, the model estimates organic carbon decomposition and derives mineralization rates from two pools representing the organic N and the added organic N (Stella, 2015; Manzoni et al., 2012). The model represents several management options such as N fertilization, irrigation and residue management, but does not include the effect of tillage. In *SiriusQuality* nitrogen can be lost from the soil either by N₂O emission, denitrification (N₂) or leaching, but volatilization is not implemented, and the model does not consider N immobilization. The crop module currently simulates wheat and barley growth.

2.3.3. *STICS*

The *STICS* model (version 9.2; Brisson et al., 2010) simulates the water, C and N fluxes with a soil vertical grid of 1 cm (description of the soil is however done at the scale of larger soil layers). Soil temperature is also simulated with the same grid of 1 cm. Organic matter turnover is simulated based on four main pools: organic residues, microbial biomass, humus and mineral pools (CO₂ and N). It includes different soil processes (mineralization, immobilization, nitrification, volatilization, denitrification, leaching) as well as source/sink effect of the crop (symbiotic N fixation, absorption of mineral N). The model can take into account a large diversity of crops and the main crop and soil management practices, including mineral and organic fertilization, irrigation, soil tillage and residues management.

2.3.4. Model coupling and calibration

The nitrification, denitrification and N₂O emission module is driven by five input variables: NH₄⁺ and NO₃⁻ concentration, soil temperature, WFPS, and pH. These inputs are either constant variables (pH), external variables, which depend only on processes external to the module (soil temperature and WFPS), or internal variables (soil NO₃⁻ and NH₄⁺ concentration) that are affected by internal processes and continually updated and exchanged with the main model (Fig. 2).

In previous studies, this N₂O module was coupled to the *STICS* model and used to simulate the N₂O emissions in oceanic and Mediterranean climate conditions (Peyrard et al., 2017; Plaza-Bonilla et al., 2017). In this study, we coupled it with the models *MONICA* and *SiriusQuality* by means of the internal and external variables (Fig. 2). The nitrification/denitrification and N₂O module was integrated directly into the source code of the models and was therefore implemented in C# (*SiriusQuality*) and C++ (*MONICA*), whereas the original version was in Fortran (*STICS*). The previous nitrification/denitrification and N₂O module of *MONICA* was disabled for this study. In *SiriusQuality*, the existing denitrification formalism was replaced by the new module.

Soil NO₃⁻ and NH₄⁺ concentrations were provided by the soil nitrogen component of each agroecosystem model and affected by the internal and external soil N fluxes, such as mineralization, fertilization, leaching, volatilization, and N uptake by roots for crop growth. Soil WFPS and soil temperature were provided by the soil water and soil temperature components, respectively. Since the models do not represent the migration of the gas in the soil subsurface, we assumed that the N₂O produced in soil was immediately released into the atmosphere, and we neglected the effect of soil moisture on tortuosity and gas diffusion (Chamindu Deepagoda et al., 2019).

The three agroecosystem modeling groups performed a minimal calibration of cultivar parameters using phenology and grain yield data from the different treatments listed in Table 1. Cultivar parameters were minimally adjusted to observed phenology using the trial-and-error method. If the same cultivar was present over several years or treatments, *SiriusQuality* used the Nelder Mead algorithm to minimize the RMSE of observed and simulated anthesis and maturity dates, as well as grain yield. Grain yield information was available for all treatments, while anthesis date was available for all treatments except for EM.19 and maturity date was available only for the experiments carried out at Ottawa, Grignon, New Delhi, and Santa Maria. Given the limited amount of information and the number of cultivars to be calibrated, we used the

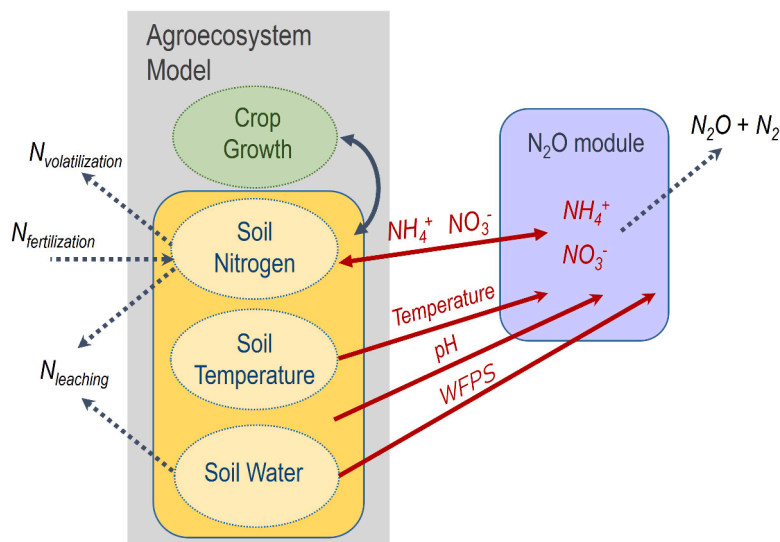


Fig. 2. Schema of the coupling of the N₂O module with the relevant model components of the agroecosystem models. Driving variables (internal and external) of the N₂O module are represented in red, solid blue arrows represent N or water fluxes within the agroecosystem model, and dotted arrows represent N fluxes that enter or are lost from the crop-soil system.

entire phenology and grain yield data for calibration.

There was no calibration of the N₂O module in this study, to ensure that all models used the same default module parameterization and that inter-model variation in N₂O emission were due to differences external to the module. For each of the 31 treatments, the models simulated the daily N₂O emission, soil NO₃⁻ and NH₄⁺ concentrations, WFPS, soil temperature, leaf area index (LAI), total above ground crop biomass and crop nitrogen (Supplementary Figs. S1–S8).

2.4. Uncertainty and sensitivity analysis

With the exception of pH, which has a constant value during the simulation of a cropping season, the value of the input variables driving the N₂O emission module change during the simulation. Their values at a given time step differ between agroecosystem models, due to the different representation of processes, different formalisms, parameters and hypotheses. The range of simulated values represents the structural model uncertainty and can be characterized by the daily coefficient of variation of variable *i* at treatment *j* and timestep *t* ($C_{Vi,t}^j$), which quantifies the daily dispersion of the variable around the ensemble mean. For a given treatment and variable, the daily $C_{Vi,t}$ values were obtained from the standard deviation σ_i of the variable simulated by the three models, divided by the ensemble mean μ_i :

$$C_{Vi,t} = \frac{\sigma_i}{\mu_i} \quad (1)$$

$$\sigma_i = \sqrt{\frac{1}{3}(x_{MO,t} - \mu_i)^2 + (x_{SQ,t} - \mu_i)^2 + (x_{ST,t} - \mu_i)^2} \quad (2)$$

$$\mu_i = \frac{1}{3}(x_{MO,t} + x_{SQ,t} + x_{ST,t}) \quad (3)$$

Where $x_{MO,t}$, $x_{SQ,t}$, and $x_{ST,t}$ are the values simulated at timestep *t* by MONICA, SiriusQuality and STICS, respectively.

For each treatment *j* and input variable *i*, we estimated the uncertainty U_i^j using the median of the $C_{Vi,t}^j$:

$$U_i^j = \text{median}(C_{Vi,t}^j) \quad (4)$$

The choice to use the median rather than the mean is motivated by the lower sensitivity of the median to outliers. Uncertainty was estimated on the 0–30 cm layer, by calculating the average over the 0–10 cm, 10–20 cm and 20–30 cm layers. The depth of 30 cm corresponds to the default value of the maximum depth of nitrification in the module, while for denitrification the default value of the maximum depth 20 cm. The average uncertainty of an input variable was estimated by calculating the mean U_i^j over all treatments *j*:

$$\bar{U}_i = \text{mean}(U_i^j) \quad (5)$$

The impact of the uncertainty of the input variables on N₂O emission was evaluated by performing a sensitivity analysis (SA) of the N₂O module with the Extended Fourier Amplitude Sensitivity Test (eFAST; Saltelli et al., 1999) implemented in the R-package Sensitivity (Iooss et al., 2022). This sensitivity analysis scheme is a robust, variance-based and computationally efficient approach for global sensitivity analysis, that can be applied for complex, non-linear models and represents the sensitivity for an individual factor and for total factor interactions (Saltelli et al., 2008). For each input (or driving variable) the original trajectory during a simulation was modified by multiplying the value by a factor, randomly chosen from a uniform distribution. We performed two types of SA. In the “fixed range” SA, the uniform distribution was constant between variables ($\pm 20\%$ of the nominal value). In the “uncertainty-driven” SA, the upper and lower bound of the distribution were set using the mean uncertainty of the variable (Eq (5)) over all treatments. This means that for the variables with larger uncertainties, a larger spectrum of trajectories was tested in the uncertainty-driven SA.

The effect on the soil N pool (NO₃⁻ and NH₄⁺) and the feedback on crop growth were not considered. The pH was not included in the SA as a variable, because the value is fixed during the simulation and is not concerned by model uncertainty, but we included its effect by performing SA for acidic (pH < 6.5) and basic (pH > 7.5) soils, separately.

2.5. Analysis of N₂O model error

Different approaches were used to analyze the sources of the model error in predicting N₂O emissions at different scales (daily time step and over the growing season). We applied the Classification and Regression Tree modeling approach (CART modeling; Breiman et al., 1984) to describe the factors that affect the model error at daily scale. This algorithm allows to create decision tree for non-categorical data and constructs a binary tree based on the minimization of the Gini Impurity. The algorithm aims at producing rules that predict the value of an outcome variable from known values of explanatory variables. The algorithm helps to explore the structure of data sets and develops a binary decision tree. Each node of the decision tree is a rule that splits the data in two groups, with the aim of maximizing homogeneity within a group. The process is then applied to each sub-group recursively until the subgroup reaches a minimum size or until no improvement can be made. Then a cross validation is performed to prune the tree and identify the tree with the lowest cross-validated error to avoid overfitting.

We generated regression trees through the Recursive Partitioning and Regression Trees approach using the R package Rpart.plot (Milborrow, 2021) with the objective of predicting the N₂O emission error (the difference between daily simulated and measured N₂O emission) based on the simulated input variables. The predictor variables were the NO₃⁻ and NH₄⁺ concentration, the soil temperature, the WFPS and the pH in the 0–30 cm soil layer.

The random forest algorithm (RFA), based on the analysis of multiple randomly created decision trees (Breiman, 2001), was also applied to identify the most important variables (R package randomForest; Liaw and Wiener 2002). We used the increase in Mean Squared Error of predictions (%IncMSE) to represent the mean decrease in accuracy when the variable is left out. The higher the value of the %IncMSE, the higher the importance of the variable to the model. The %IncMSE was calculated as:

$$\%IncMSE = \frac{\sum_{ntree} MSE_{0,i} - MSE_{J,i}}{ntree} \frac{1}{\sigma_{MSE_{0,i} - MSE_{J,i}}} \quad (6)$$

Where $MSE_{0,i}$ is the mean squared error of the regression, $MSE_{J,i}$ is the mean squared error of the regression when the variable *J* is permuted, *i* is the index of the tree, *ntree* is the total number of generated trees, and $\sigma_{MSE_{0,i} - MSE_{J,i}}$ is the standard deviation of the differences between the permuted and not permuted *MSE* of regressions.

Decision trees are very easy to interpret because they allow to visualize the rules (variables and values) of the splits. However, their result is less stable than that of the RFA and depends on the training set. RFA generates many trees, then averages all the trees, and therefore produces a more stable and accurate result than simple decision trees. However, RFA does not allow visualization of the final model, which limits interpretation. By combining the CART and RFA algorithms we obtain a robust result and information on the rules, which helps to better understand the conditions associated with a larger error.

The model error (*ME*) was also analyzed over the growing season, and we compared the simulated cumulative N₂O emission to the measured N₂O emission from sowing to harvest. To better compare errors between experiments, we represented the model error as the absolute value of the relative error:

$$ME = \frac{\left| \sum_{\text{sowing}}^{\text{harvest}} N_2O_{\text{obs}} - \sum_{\text{sowing}}^{\text{harvest}} N_2O_{\text{sim}} \right|}{\sum_{\text{sowing}}^{\text{harvest}} N_2O_{\text{obs}}} \quad (7)$$

where N_2O_{obs} ($kg\ N_2O-N\ ha^{-1}$) is the measured N_2O emission, and N_2O_{sim} ($kg\ N_2O-N\ ha^{-1}$) is the simulated N_2O emission.

In some experiments the time series of N_2O measurement were not complete and the missing data were estimated using a spline interpolation.

All analyses were performed using the R statistical software program version 4.0.5 (R Core Team, 2021).

3. Results

3.1. Model evaluation for final total above ground biomass and nitrogen, and grain yield and nitrogen

Overall model performance was evaluated by comparing simulated and observed values of selected variables. The evaluation was done for each of the agroecosystem models and for the multi-model ensemble average (e.mean).

3.1.1. Final total above ground biomass and nitrogen, and grain yield and nitrogen

We compared simulated and observed values of final total above ground N, grain N, grain yield, and final total above ground biomass (Fig. 3). The best performance of e.mean was found for final total above ground N, followed by grain yield. *SiriusQuality* performed better than the e.mean for grain N and final total above ground mass. The worst model prediction was obtained by MONICA for grain N (RRMSE = 38%), but the overall performance of the models was satisfactory, with a RRMSE between 14% and 29%.

3.1.2. Ammonium, nitrate, WFPS and soil temperature

We compared the simulated and measured value of the four N_2O module input variables in the first soil layer: ammonium, nitrate, WFPS and soil temperature (Fig. 4). The depth of the first soil layer is 30 cm for the CA and EM treatments, and 10 cm for all other treatments. Soil temperature showed the best accuracy, with RRMSE varying between 9% (e.mean) and 18% (*SiriusQuality*). The second best simulated input

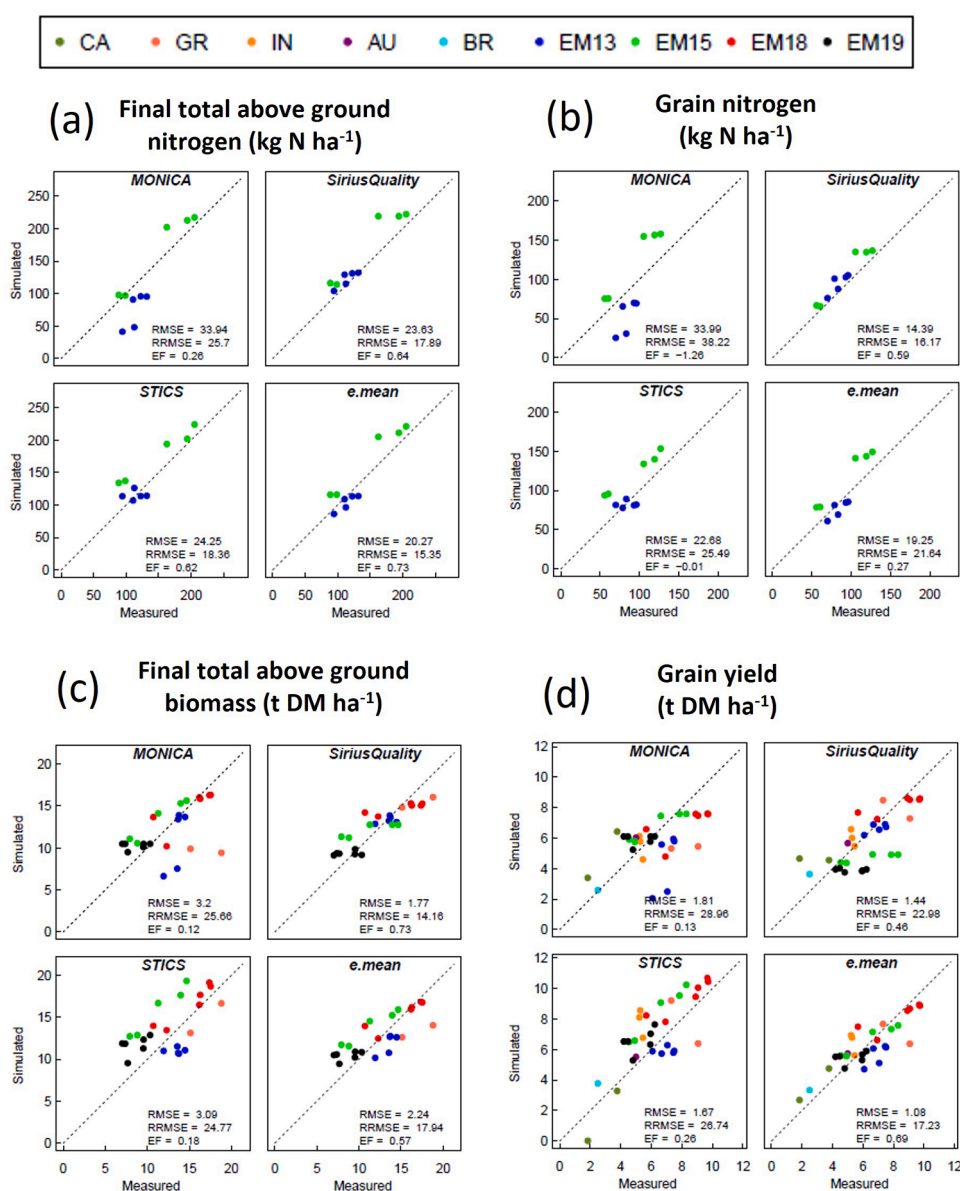


Fig. 3. Comparison of simulated and measured variables for the three agroecosystem models (*MONICA*, *SiriusQuality* and *STICS*) and for the ensemble mean (e.mean). (a) Final total above ground nitrogen, (b) grain nitrogen, (c) grain yield, and (d) final total above ground biomass. RMSE and RRMSE values are given at the lower left corner of each panel. The colors of the symbols represent the different locations of the experiments (or locations/year combination for Estrées-Mons).

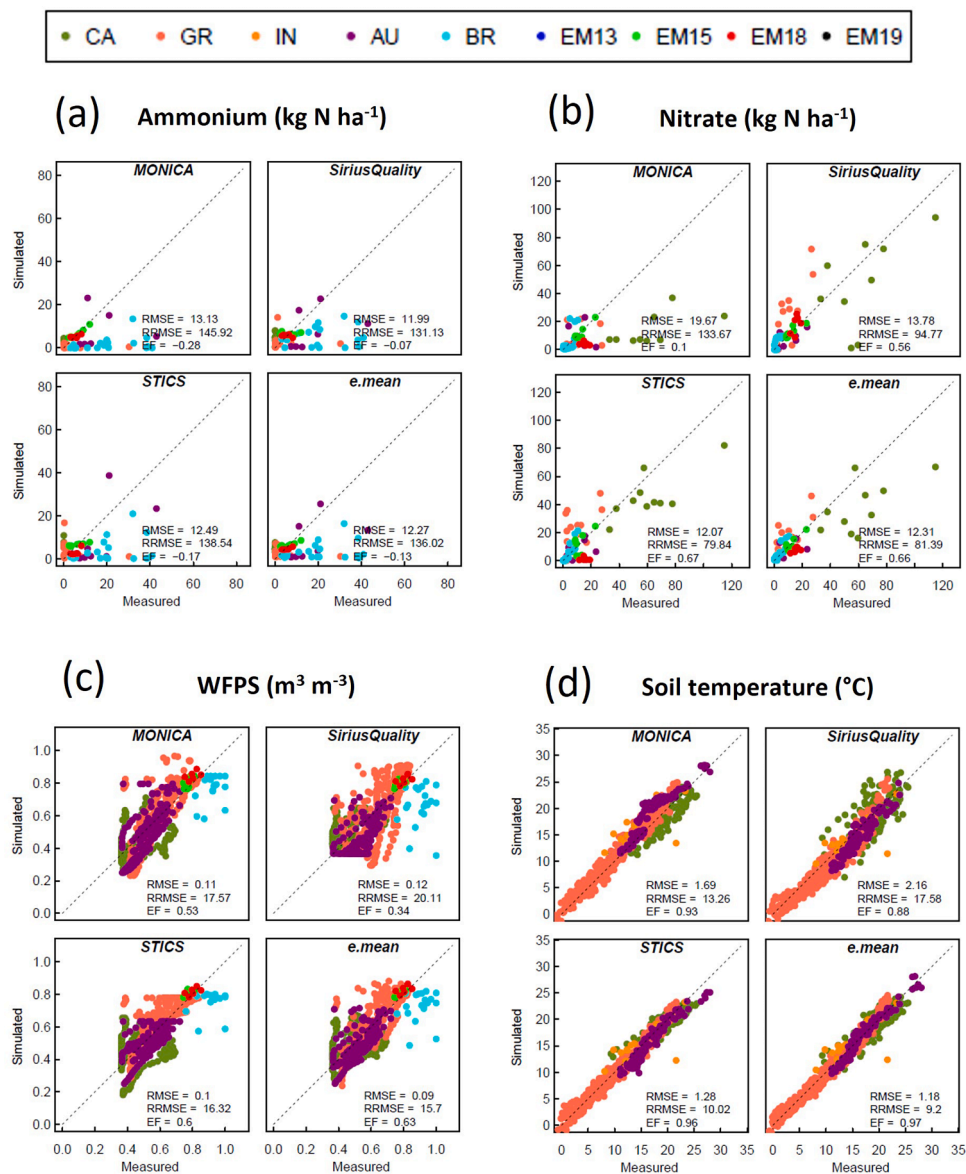


Fig. 4. Comparison of simulated and measured variables for the three agroecosystem models (MONICA, *SiriusQuality* and STICS) and for the ensemble mean (e.mean) in the first soil layer. (a) Ammonium, (b) nitrate, (c) WFPS, and (d) soil temperature. RMSE and RRMSE values are given at the lower left corner of each panel. The colors of the symbols represent the different locations of the experiments (or locations/year combination for Estrées-Mons).

variable was WFPS, with RRMSE between 15% (e.mean) and 21% (*SiriusQuality*). The accuracy of nitrate and ammonium was poor. For nitrate the models showed RRMSE varying between 80% (STICS) and

134% (MONICA), and for ammonium between 131% (*SiriusQuality*) and 146% (MONICA). The simulation of ammonium at the Brazilian site was particularly poor for all models and showed large discrepancies between

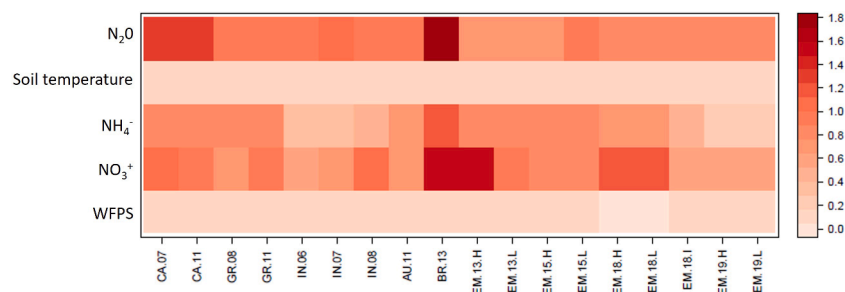


Fig. 5. Heat map of the uncertainty of simulated N_2O emission and the four input variables of the N_2O module, represented by the uncertainty estimator U_i^j for the agroecosystem models MONICA, *SiriusQuality*, and STICS for the field experiments considered in this study. Simulated water filled pore space (WFPS), NO_3^- content, NH_4^{++} content, and soil temperature were average over the 0 to 30 cm soil layers. The treatments are defined in Table 1.

simulated and measured values. Simulation accuracy of daily and cumulative N_2O emissions was also low, particularly for the Canadian and Brazilian experiments. (Fig. S9).

3.2. Uncertainty of the N_2O module variables

The uncertainty of the four N_2O module input variables (WFPS, NO_3^- , NH_4^+ , and soil temperature) and of the output variable (N_2O emissions) was estimated for each treatment (Eq (4)). WFPS and soil temperature had the lowest uncertainty, while NO_3^- and NH_4^+ had the highest (Fig. 5). The uncertainty of N_2O emissions was also high, especially for the Canadian and Brazilian experiments. The mean uncertainty (Eq (5)) was also estimated, and for WFPS and soil temperature we obtained values of 0.10 and 0.11, respectively, while for NO_3^- and NH_4^+ we estimated values of 0.91 and 0.67, respectively.

3.3. Sensitivity analysis of the N_2O module

The uncertainty-driven SA, that used the average uncertainties of the input variables to set up the range of values to be tested, was compared to the fixed range SA, which does not take into account prior knowledge of the uncertainty of the input variables and assumes the same value for all tested variables ($\pm 20\%$ of the nominal value). The results showed large differences between the uncertainty-driven and the fixed range SA, and between acidic and basic soil. In the uncertainty-driven SA, N_2O emission of basic soils were mainly impacted by NH_4^+ concentration, while for acidic soils NO_3^- was the most sensitive input variable, followed by WFPS (Fig. 6). In the fixed range SA, the most sensitive parameter was soil temperature, followed by NH_4^+ concentration and WFPS in basic soils, and WFPS, followed by soil temperature in acidic soils. The analysis highlighted that the large uncertainty characterizing the simulation of NH_4^+ and NO_3^- translates into a high sensitivity of these inputs in the SA, while WFPS and temperature have a lower variability and therefore a lower sensitivity.

3.4. Pedo-climatic conditions and N_2O emission error

Decision trees were developed using the CART algorithm, to predict N_2O emission error (Eq (7)) based on modelled pedo-climatic conditions (WFPS, pH, NO_3^- , NH_4^+ and soil temperature) for each agroecosystem model and for the e.mean (Fig. 7). At the same time, the RFA was used to rank the variables according to their importance to the model error. In

SiriusQuality and e.mean the most important variable to determine model error according to RFA was pH and in *SiriusQuality* the highest error values were found in acidic soils and high NO_3^- concentrations. In the STICS model, the effect of soil pH was not dominant, but the most important variable to determine model error was the NO_3^- concentration, followed very closely by WFPS, and the largest error occurred with high NO_3^- concentration and wet soils. The MONICA model gave a less clear-cut result, with two variables of similar importance for model error: NH_4^+ and pH. Similar to *SiriusQuality*, the largest error occurred in acidic soils with high NO_3^- concentration. Although the graphs and rankings varied between models, we observed that pH and soil N concentration were the most important variables to determine the model error.

3.5. Relative error for seasonal cumulative N_2O emission

The relative error for seasonal cumulative N_2O emission was calculated for each combination of site/year/N treatment, from sowing to harvest (Fig. 8). The results highlighted three situations in which at least one agroecosystem model had low accuracy: at Santa Maria, Brazil (BR13), for the three models, at Ottawa, Canada, in 2011 (CA.11) for the STICS model, and in some treatments at Estrées-Mons, France, for which both MONICA and STICS showed large error, in particular during the winter wheat experiment 2014–2015 (EM.15.L). In the following section, we analyze in more detail the factors that caused high model error in these treatments

3.5.1. Estrées-Mons, France (EM.15.L)

In this treatment, the winter wheat cultivar Cellule was sown in a basic soil (pH 8.2). All three agroecosystem models agreed that an initial peak in N_2O emissions occurred at sowing due to incorporation of crop residue, although mineralization of residues leads to different soil NH_4^+ increases between the models (Fig. 9). A second peak was measured after the N fertilizer application (day after sowing [DAS] 164), but the height and timing of the peak differed among the models. In *SiriusQuality*, the N_2O peak occurred several days later than the date of application because the model has a condition of minimum soil moisture and amount of rainfall in the days before fertilization for the fertilizer to penetrate the soil.

The simulated N_2O emissions from nitrification and denitrification were in the same range. However, the denitrification rates diverged between the models. All three models have the same potential

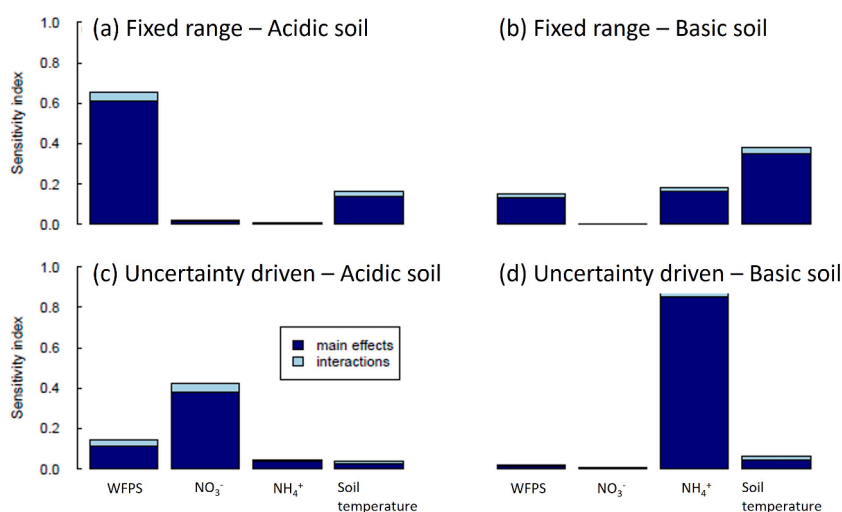


Fig. 6. Sensitivity analysis (SA) of the N_2O module. (a) and (b) fixed range SA assuming that all variables had the same uncertainty for acidic (pH < 6.5) and basic (pH > 7.5) soils, respectively. (c) and (d) uncertainty-driven SA performed using the average multi-model ensemble uncertainty of each input variable for acidic and basic soils, respectively.

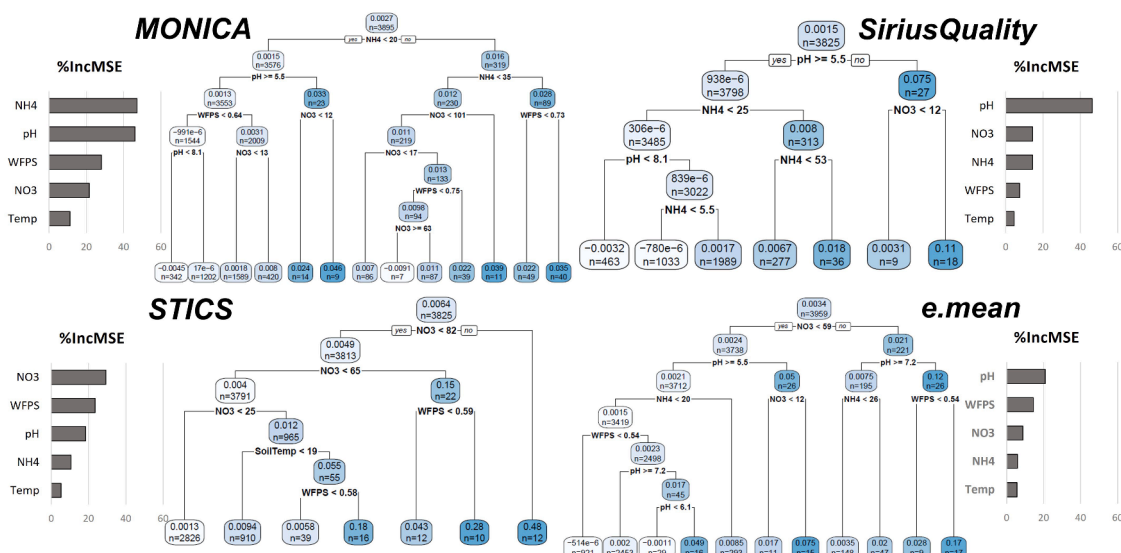


Fig. 7. Classification and regression tree (CART) of the daily difference between simulated and measured N_2O emissions (model error) for the agroecosystem models MONICA, *SiriusQuality* and STICS, and the multi-model ensemble mean (e.mean). Split point of the decision tree represent constrains to make the prediction. Nodes show the predicted values and the percentage of measurements in the node. Darker nodes represent larger errors. Bar plots represent the output of the Random Forest analysis and show the mean decrease of accuracy, with higher values corresponding to higher importance of the variable in the model.

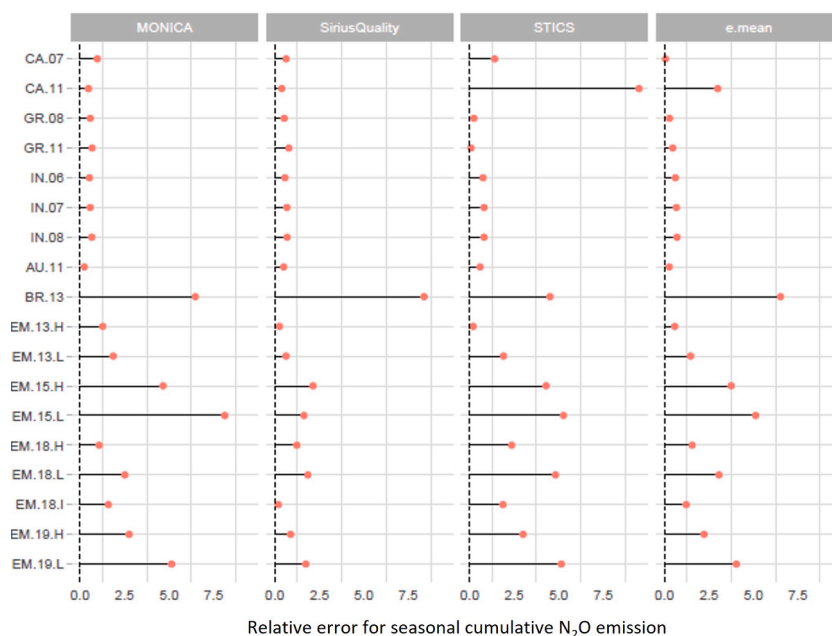


Fig. 8. Relative error for seasonal cumulative N_2O emissions from sowing to harvest for the field experiments consider in this study simulated with the agroecosystem growth models MONICA, *SiriusQuality* and STICS, and the multi-model ensemble mean (e.mean). For Estrées-Mons, treatments were grouped by the rate of N fertilizer applications (L: low, I: intermediate, H: high). The treatments are defined in Table 1.

denitrification rate, but the weighting factors, which reduce the potential denitrification as a function of WFPS, temperature and NO_3^- values, depend on the model. After DAS 150, *SiriusQuality*'s WFPS weighting factor is close to zero and strongly limits denitrification, whereas MONICA and STICS show peaks of denitrification rate linked to WFPS fluctuations (Fig. S10). This shows that water retention and water use differed between models, with a significant impact on denitrification and N_2O emissions. In addition, *SiriusQuality* had a smaller error (lower overestimation) compared to the two other models also because the NO_3^- and NH_4^+ peak after fertilization was smaller.

3.5.2. Santa Maria, Brazil (BR.13)

In this experiment a spring wheat was sown in an acidic soil (pH 5.4). The experiment is characterized by a high cumulative precipitation during the growing season, the highest among the experimental sites where spring wheat was sown.

In the simulations, the main source of N_2O emissions was denitrification and all models significantly overestimated these emissions (Fig. 10). The experiment had no nitrogen fertilization, but the initial soil inorganic nitrogen was high (37 kg N ha^{-1} for the 0–10 cm layer) and crop residues (corn) were left on the field. As a result, the models simulated an initial peak in NO_3^- and NH_4^+ , and high initial denitrification. Again, the amplitude and the duration of the peak was fairly

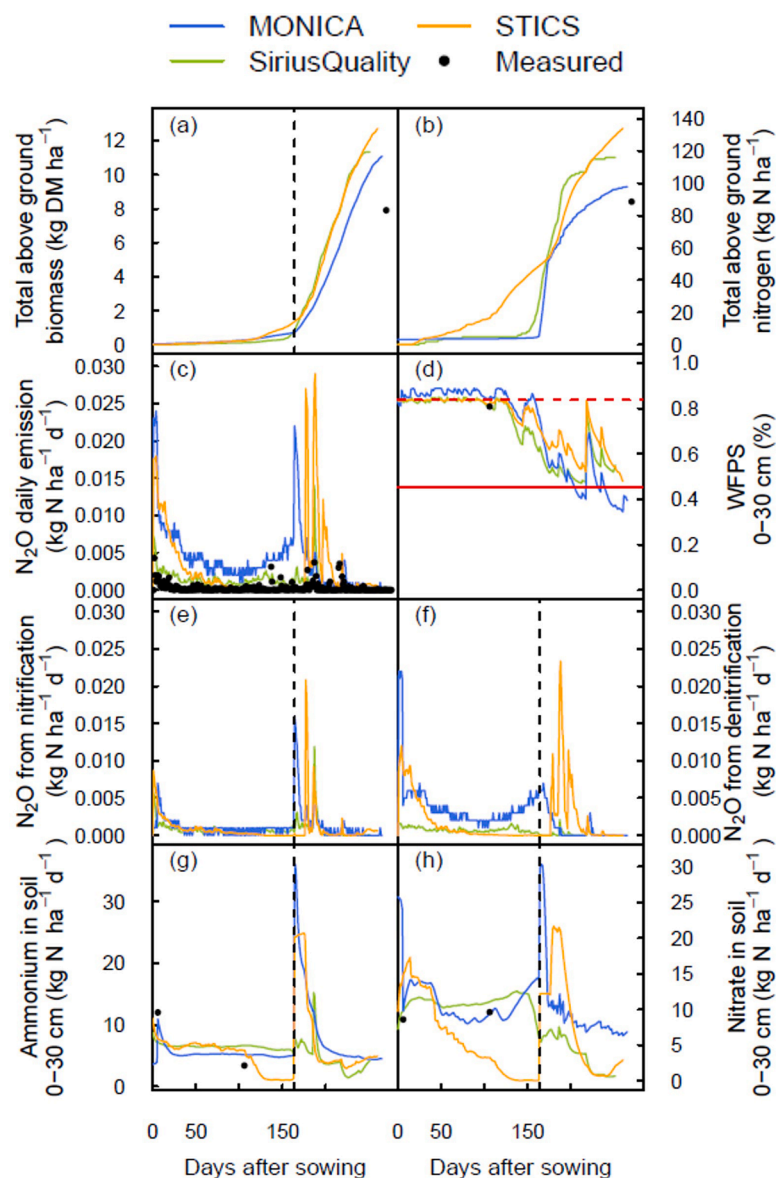


Fig. 9. Simulated (lines) and measured (circles) variables versus days after sowing for the winter wheat cultivar Cellule grown in the field at Estrées-Mons, France during the 2015–2016 growing season with low N supply (treatment EM15.L). (a) Total above ground crop biomass. (b) Total above ground nitrogen. (c) Daily N_2O emission from the top 0–40 cm soil layer. (d) Daily N_2O emission from nitrification from the top 0–40 cm soil layer. (e) Daily N_2O emission from denitrification from the top 0–20 cm soil layer. (f) Soil NH_4^+ content in the top 0–30 cm soil layer. (g) Soil NO_3^- content in the top 0–30 cm soil layer. In (d), dotted and solid horizontal red lines show field capacity and permanent wilting point, respectively.

variable between the three models.

After DAS 70, MONICA simulated higher soil NO_3^- concentrations in the 0–10 cm soil layer than the other models and thus simulated a higher N_2O emission due to denitrification. This is favored by a smaller decrease in WFPS, which allows for conditions suitable for denitrification. It is interesting to note that in the measurements, WFPS were often higher than the field capacity, indicating very wet conditions, which *SiriusQuality* and *STICS* were not able to simulate, as simulated excess water moved quickly through the soil profile. This is likely an important factor that is misrepresented in the models and introduces errors in the simulation of not only N_2O emission but likely also of crop growth to waterlogging conditions.

3.5.3. Ottawa, -Canada (CA.11)

In the Canadian treatment, a spring wheat (ACBrio) was sown in a slightly acidic soil (pH 6.6). There was one fertilization event (with urea), just before sowing. The models simulated an initial peak of NH_4^+ , followed by a later peak of NO_3^- . In this experiment, there was a large difference between the *STICS* simulations and those of *MONICA* and *SiriusQuality*, and *STICS* had the largest error, due to the overestimation of denitrification and N_2O emission (Fig. 11).

In this case, the difference between denitrification rates was too great to be attributed to weighting factors alone. In *STICS*, denitrification followed the same fluctuations as the WFPS weighting factor, the main limiting factor (Fig. S11). However, the denitrification rate estimated by multiplying the potential denitrification with the weighting factors for temperature, WFPS and NO_3^- simulated in the 0–10 cm and 10–20 cm layers, was much lower. As the vertical grid of the *STICS* model was 1 cm, whereas in *MONICA* and *SiriusQuality* the variables were averaged over thicker layers (10 cm and 5 cm, respectively), we tested whether a variation in WFPS in the 0–20 cm soil layer could explain the large denitrification simulated by *STICS*. Since the *STICS* model does not allow to retrieve the WFPS value per 1 cm layer, we implemented the denitrification formalism in an R environment. Noise was applied to the WFPS values of the 0–10 cm and 10–20 cm layers simulated by *STICS* in order to get plausible WFPS values for each 1 cm layer. The denitrification rate was then estimated for each 1 cm layer and summed to obtain total denitrification (Fig. S12). This test allowed to produce the range of denitrification rate values simulated by *STICS* for this experiment, and showed that the finer vertical grid, combined with the strong nonlinear response of denitrification and N_2O emission to WFPS, were a plausible cause of the large overestimation of

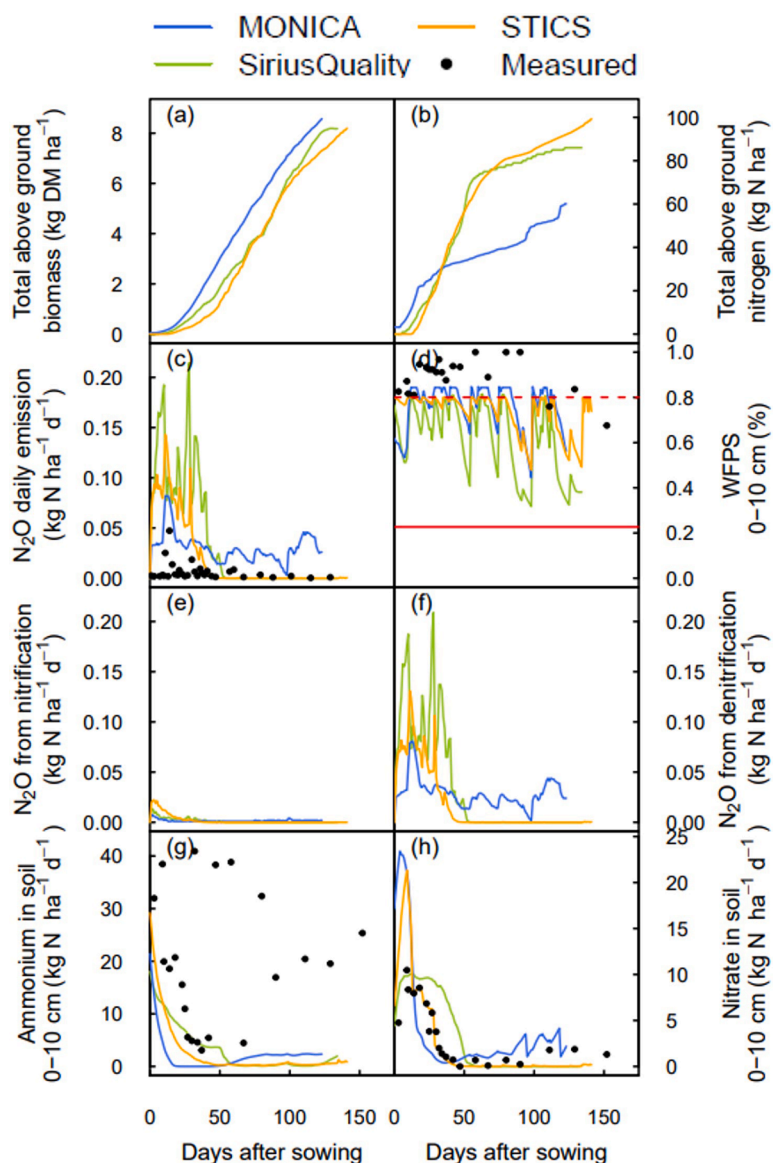


Fig. 10. Simulated (lines) and measured (dots) variables versus days after sowing for the spring wheat cultivar Quartzo grown in the field at Santa Maria, Brazil, during the 2013 growing season. (a) Total above ground crop biomass. (b) Total above ground nitrogen. (c) Daily N_2O emission from the top 0–40 cm soil layer. (d) Water filled pore space in the 0–30 cm soil layer. (e) Daily N_2O emission from nitrification from the top 0–20 cm soil layer. (f) Daily N_2O emission from denitrification from the top 0–20 cm soil layer. (g) Soil NH_4^+ content in the top 0–30 cm soil layer. (h) Soil NO_3^- content in the top 0–30 cm soil layer. In (d), dotted and solid horizontal red lines show field capacity and permanent wilting point, respectively.

denitrification by STICS.

4. Discussion

The growing effort toward modularity in agricultural and ecological models is a driving force for improvement of these models because it allows easy implementation and exchange of modules and testing of assumptions and formalisms. However, when the same module is implemented in different models, the outcome depends not only on the appropriateness of the formalisms used to represent the biophysical processes, but also on the output of other connected modules that provide the input variables for the module. Therefore, it is necessary to study the link between the module and the model in which it is integrated, and to assess how the uncertainty in driving variables affects the uncertainty of the model prediction. The novelty of this study lies in the effort to evaluate the performance of a module when it is coupled to different agroecosystem models, with the objective of elucidating the causes and conditions of greater or lesser model accuracy.

We evaluated the variability of the simulated N_2O emissions, and characterized the model performance, without calibration of the potential denitrification rate. Indeed our objective was not to evaluate the module, as this had already been done in previous studies (Peyrard et al.,

2017; Plaza-Bonilla et al., 2017). Instead, we were interested in the range of outputs that can be obtained by integrating an uncalibrated module (with default parametrisation) into several agroecosystem models. This allowed us to assess the predictive ability of the module without calibration as well as the variability in accuracy between agroecosystem models in different environments, and the reasons for this variability. There is a need to quantify N_2O emissions on a large scale and therefore to apply models that are sufficiently robust and general to accurately predict N_2O emissions in a wide range of pedo-climatic conditions. Currently, most of the countries estimate national N_2O emissions using IPCC Tier 1 approach, based on a global default emission factor (IPCC, 2006), and some countries have moved to country-specific emission factors (Tier 2). However, additional efforts are needed to implement model-based Tier 3 method, and agroecosystem models provide a structured way to better account for complex interactions between soil properties, weather, and crop management practices that affect N_2O emission (Brilli et al., 2017; Fuchs et al., 2020; Khalil et al., 2016).

Characterization of the simulation uncertainty identified the greatest variability in NO_3^- and NH_4^+ concentration, while WFPS and soil temperature showed fairly consistent values among the different agroecosystem models. In addition, the model evaluation showed that soil

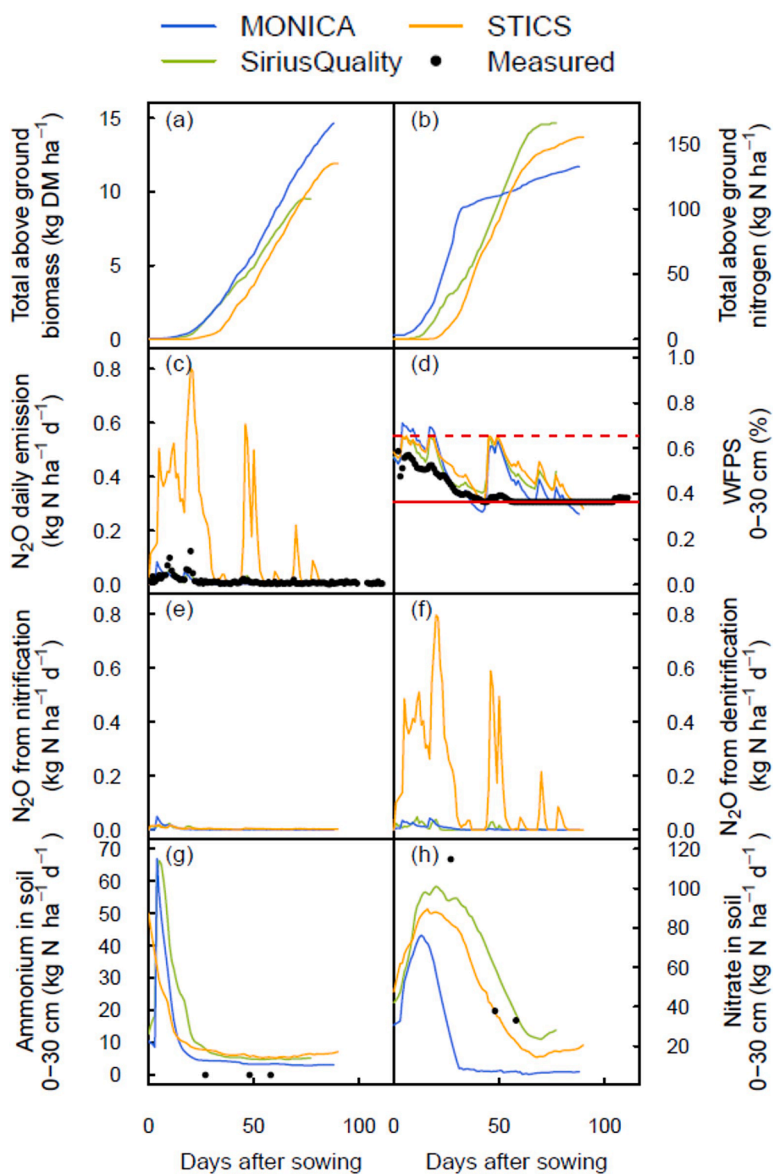


Fig. 11. Simulated (lines) and measured (dots) variables versus days after sowing for spring wheat ACBrio grown at Ottawa, Canada, during the 2011 growing season. (a) Total above ground crop biomass. (b) Total above ground nitrogen. (c) Daily N_2O emission from the top 0–40 cm soil layer. (d) Water filled pore space in the 0–30 cm soil layer. (e) Daily N_2O emission from nitrification from the top 0–40 cm soil layer. (f) Daily N_2O emission from denitrification from the top 0–20 cm soil layer. (g) Soil NH_4^+ content in the top 0–30 cm soil layer. (h) Soil NO_3^- content in the top 0–30 cm soil layer. In (d), dotted and solid horizontal red lines show field capacity and permanent wilting point, respectively.

temperature and WFPS were simulated more accurately than NO_3^- and NH_4^+ . Previous studies had already highlighted that the simulation of soil inorganic nitrogen, NH_4^+ and NO_3^- was generally less accurate than the simulation of soil water and temperature (Coucheny et al., 2015; Smith et al., 2019). The processes involved in the nitrogen cycle are complex and depend on biotic and abiotic conditions, which are spatially variable even on a fine scale (Liang et al., 2023; Yin et al., 2020). Consequently, high spatial variability is often observed, particularly for NH_4^+ and NO_3^- . Model uncertainty and larger measurement errors lead to lower model performance indicators (Coucheny et al., 2015). Furthermore, recent studies have shown that in (sub)tropical agroecosystems, the activity and relative importance of microbial communities involved in nitrification are different from those in temperate soils (Zhang et al., 2023; Fan et al., 2021). The ratio of nitrate to ammonium is also lower in tropical soils than in temperate soils, but current agroecosystem models are not yet able to simulate these differences.

The uncertainty of N_2O emissions was also high, especially for the Canadian and Brazilian experiments, which showed the largest model error for cumulative and daily N_2O emission. According to our results, the variables with the highest uncertainties were also the less accurate. Our characterization of uncertainty of input variables was used to

characterize the sensitivity of the N_2O module, and our results showed that the variables with the greatest impact on the simulation of N_2O emission were NH_4^+ and NO_3^- concentration. It also showed that the sensitivities were strongly dependent on soil pH.

Previous research has found that good model performance in representing cumulative N_2O emissions at the annual scale does not always coincide with good representation of daily emissions (Fuchs et al., 2020) and that it is interesting and complementary to study both. We therefore examined the model error at the daily scale by calculating the difference between measured and simulated daily N_2O emission, and at the “growing season” scale, by comparing cumulative N_2O emission from sowing to harvest. The Random Forest and CART algorithm showed that pH was among the main factors of daily model error for *SiriusQuality*, *MONICA* and *e.mean*, while for *STICS* the main factor was NO_3^- . The highest model error was found when simulating acidic soils with high NO_3^- concentration in *SiriusQuality*, *MONICA* and *e.mean*, and when simulating high NO_3^- concentration and wet soils in *STICS*. These conditions, which are challenging to simulate accurately, are favorable to denitrification. However, the analysis of model error is based on simulated input variables, which limits the inference that can be drawn from the results. We do not evaluate how real conditions affect model error,

but only how simulated input variables affect model error. If a model has a systematic bias in the representation of an input variable and underestimates a variable, and if the model is sensitive to this variable, the regression trees may fail to detect the factor, because the variable is poorly simulated.

To characterize the performance of the models over the growing season we examined the evolution of several variables from sowing to harvest. We analyzed the external factors that affect the value of the variable, such as crop management and weather, and we analyzed the relationship between the model error and the shift in the peaks of the driving variables. The analysis showed that all models simulated NO_3^- and NH_4^+ peaks after fertilization and crop residue incorporation, but the timing and magnitude of the simulated peaks could vary considerably between models, with a major effect on N_2O emission. After a nitrogen fertilizer application, we observed a lag between simulated peaks of several days, for example due to soil moisture or rainfall conditions that are required in *SiriusQuality* for the fertilizer to penetrate the soil surface. Similarly, the three agroecosystem models represent the incorporation and mineralization of crop residues with different formalisms and parameters, resulting in N peaks with different magnitude and shape. Another factor that could have a significant impact on the model results is the vertical grid used to simulate the soil layers in the models. The STICS model that had the finest vertical soil grid was also the model that most overestimated denitrification in the second year of the Canadian experiment. Since nitrification and denitrification are affected by non-linear processes, small differences in the representation of soil conditions or soil layers have substantial implications for the simulation of N_2O emissions by the module. Finally, there was a general trend of overestimating N_2O emissions from denitrification in all three models, which could be corrected by a calibration of the denitrification potential.

Another important achievement of our study is the evaluation of N_2O emission module in a variety of contrasting soil-climate environments. Our analysis has emphasized that wet acidic soils with high denitrification potential were more challenging for models to simulate. Since soil pH is a key factor for estimating regional N_2O emission (Wang et al., 2018) this limitation has important consequences on model robustness and currently limits the use of process-based models to estimate N_2O emissions at the regional scale. Agroecosystem models tend to be less successful at sites to which they have not been previously calibrated (Brilli et al., 2017) and wet acidic soils are unusual conditions for them. Around 75% of denitrification studies were located in Europe and North America and the underrepresentation of tropical and sub-tropical agricultural soils is a major limitation for model validation (Almarazetal et al., 2020; Albanito et al., 2017). To improve our models, we need to expand the spatial coverage of N_2O emission modeling studies and to include more data on tropical regions and on acidic soils, as these are more sensitive to changes in N fertilization than alkaline soils (Wang et al., 2018) and exhibit very different nitrogen dynamics in general, as compared to soils in temperate environments, due to their iron- and aluminum-dominated mineralogy (Nendel et al., 2019).

Recent studies have suggested that the combination of process-based and machine learning (ML) models could improve the prediction of N_2O emissions (Berardi et al., 2020; Saha et al., 2021). However, to be effective, ML models must be trained on large datasets representative of the range of conditions over which the model will be used, covering different soil-climate environments and management options. Therefore, this approach has to face similar issues regarding data availability. In addition, the unbalanced structure of the data, characterized by large but sporadic N_2O peaks, is challenging for ML models, as learning the critical variables of rare events is a real issue for ML models.

5. Conclusion

In conclusion, this study has highlighted several factors that increase

uncertainty in predicting N_2O emissions, and may limit the application of models in certain soil-climate environments. Some of these factors are external to the N_2O module and depend on the representation of management and soil processes, and soil layers in agroecosystem models. For example, different formalisms for representing residue incorporation and mineralization, soil nitrogen penetration after fertilization, and soil water retention and percolation have an important impact on the variability of simulated NO_3^- , NH_4^+ , and WFPS. Other factors are internal and depend on the representation of nitrification, denitrification and N_2O emissions (e.g. the potential denitrification rate) in particular in acidic environments. We believe that greater availability of data from these environments could improve the predictive quality of the models.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors acknowledge Dr. Elizabeth Pattey (Agriculture and Agri-Food Canada, Canada), Dr. Patricia Laville (INRAE, France), Dr. Raia Silvia Massad (INRAE, France), Dr. Arti Bhatia (Indian Agricultural Research Institute, India), Dr. Massimiliano D.A. Migliorati (Queensland University of Technology, Australia) and Prof. Sandro J. Giacomini (Federal University of Santa Maria, Brazil) for granting access to the experimental data.

Funding information

This work was supported by the EIT Climate-KIC through the Barley-IT project (grant no. 1.2.6 CSAb). S.D. and P.M. acknowledge support from the Metaprogram on Agriculture and Forestry in the Face of Climate Change: Adaptation and Mitigation (CLIMAE) of the French National Research Institute for Agriculture, Food and Environment (INRAE). It also benefited from the French government grant "Investissements d'avenir", under the reference ANR-11-INBS-0001 AnaEE France, which funded the ACBB long term experiment.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.agrformet.2023.109619.

References

- Abrahamsen, P., Hansen, S., 2000. Daisy: an open soil-crop-atmosphere system model. *Environ. Mod. Software* 15, 313–330. [https://doi.org/10.1016/S1364-8152\(00\)00003-7](https://doi.org/10.1016/S1364-8152(00)00003-7).
- Aita, C., Chantigny, M.H., Gonzatto, R., Miola, E.C.C., Rochette, P., Pujol, S.B., Dos Santos, D.B., Giacomini, D.A., Giacomini, S.J., 2019. Winter-season gaseous nitrogen emissions in subtropical climate: impacts of pig slurry injection and nitrification inhibitor. *J. Environ. Qual.* 48, 1414–1426. <https://doi.org/10.2134/jeq2018.04.0137>.
- Aita, C., Schirmann, J., Pujol, S.B., Giacomini, S.J., Rochette, P., Angers, D.A., Chantigny, M.H., Gonzatto, R., Giacomini, D.A., Doneda, A., 2015. Reducing nitrous oxide emissions from a maize-wheat sequence by decreasing soil nitrate concentration: effects of split application of pig slurry and dicyandiamide. *Eur. J. Soil Sci.* 66, 359–368. <https://doi.org/10.1111/ejss.12181>.
- Albanito, F., Lebender, U., Cornulier, T., Sapkota, T.B., Brentrup, F., Stirling, C., Hillier, J., 2017. Direct nitrous oxide emissions from tropical and sub-tropical agricultural systems - A review and modelling of emission factors. *Sci. Rep.* 10 (7), 44235. <https://doi.org/10.1038/srep44235>.

- Muller, B., Martre, P., 2019. Plant and crop simulation models: powerful tools to link physiology, genetics, and phenomics. *J. Exp. Bot.* 70 (9), 2339–2344. <https://doi.org/10.1093/jxb/erz175>.
- Nendel, C., Berg, M., Kersebaum, K.C., Mirschel, W., Specka, X., Wegehenkel, M., Wenkel, K.O., Wieland, R., 2011. The MONICA model: testing predictability for crop growth, soil moisture and nitrogen dynamics. *Ecol. Model.* 222, 1614–1625.
- Nendel, C., Kersebaum, K.C., Mirschel, W., Wenkel, K.O., 2014. Testing farm management options as a climate change adaptation strategy using the MONICA model. *Eur. J. Agron.* 52, 47–56.
- Nendel, C., Melzer, D., Thorburn, P.J., 2019. The nitrogen nutrition potential of arable soils. *Sci. Rep.* 9, 5851. <https://doi.org/10.1038/s41598-019-42274-y>. Article.
- Nendel, C., Reckling, M., Debaeke, P., Schulz, S., Berg-Mohnicke, M., Constantin, J., Fronzek, S., Jakšić, S., Kersebaum, K.C., Klimek-Kopyra, A., Raynal, H., Schoving, C., Stella, T., Battisti, R., 2023. Area expansion outweighs increasing drought risk for soybean in Europe. *Glob. Chang. Biol.* 29 (5), 1340–1358. <https://doi.org/10.1111/gcb.16562>.
- Peyrard, C., Ferchaud, F., Mary, B., Gréhan, E., Léonard, J., 2017. Management practices of *Miscanthus × giganteus* strongly influence soil properties and N₂O emissions over the long term. *Bioenergy Res* 10, 208–224. <https://doi.org/10.1007/s12155-016-9796-1>.
- Plaza-Bonilla, D., Léonard, J., Peyrard, C., Mary, B., Justes, É., 2017. Precipitation gradient and crop management affect N₂O emissions: simulation of mitigation strategies in rainfed Mediterranean conditions. *Agric. Ecosyst. Environ.* 238, 89–103. <https://doi.org/10.1016/j.agee.2016.06.003>.
- Core Team, R., 2021. R: A language and Environment For Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Ravishankara, A.R., Daniel, J.S., Portmann, R.W., 2009. Nitrous oxide (N₂O): the dominant ozone-depleting substance emitted in the 21st century. *Science* 326 (5949), 123–125. <https://doi.org/10.1126/science.1176985>. Epub 2009 Aug 27. PMID: 19713491.
- Reynolds, J.F., Acock, B., 1997. Modularity and genericness in plant and ecosystem models. *Ecol. Model.* 94 (1), 7–16. [https://doi.org/10.1016/S0304-3800\(96\)01924-2](https://doi.org/10.1016/S0304-3800(96)01924-2).
- Rizzoli, A.E., Donatelli, M., Athanasiadis, I.N., Villa, F., Huber, D., 2008. Semantic links in integrated modelling frameworks. *Math. Comput. Simul.* 78 (2–3), 412–423. <https://doi.org/10.1016/j.matcom.2008.01.017>.
- Rochester, I.J., 2003. Estimating nitrous oxide emissions from flood-irrigated alkaline grey clays. *Austr. J. Soil Res.* 41 (2), 197–206. <https://doi.org/10.1071/SR02068>.
- Rötter, R.P., Palosuo, T., Kersebaum, K.C., Angulo, C., Bindi, M., Ewert, F., Ferrise, R., Hlavinka, P., Moriondo, M., Nendel, C., Olesen, J.E., Patil, R., Ruget, F., Takáč, J., Trnka, M., 2012. Simulation of spring barley yield in different climatic zones of Northern and Central Europe – A comparison of nine crop models. *Field Crop. Res.* 133, 23–36. <https://doi.org/10.1016/j.fcr.2012.03.016>.
- Saha, D., Basso, B., Robertson, G.P., 2021. Machine learning improves predictions of agricultural nitrous oxide (N₂O) emissions from intensively managed cropping systems. *Environ. Res. Lett.* 16 (2), 024004 <https://doi.org/10.1088/1748-9326/abd2f3>.
- Saltelli, A., Tarantola, S., Chan, K.P.-S., 1999. A Quantitative ModelIndependent Method for Global Sensitivity Analysis of Model Output. *Technometrics* 41 (1), 39–56. <https://doi.org/10.1080/00401706.1999.10485594>.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. *Global Sensitivity analysis: the Primer*. John Wiley & Sons.
- Sansoulet, J., Pattey, E., Kröbel, R., Grant, B., Smith, W., Jégo, G., Desjardins, R.L., Tremblay, N., Tremblay, G., 2014. Comparing the performance of the STICS, DNDC, and DayCent models for predicting N uptake and biomass of spring wheat in Eastern Canada. *Field Crop. Res.* 156, 135–150. <https://doi.org/10.1016/j.fcr.2013.11.010>.
- Smith, C.J., MacDonald, B.C.T., Xing, H., Denmead, O.T., Wang, E., McLachlan, G., Tuomi, S., Turner, D., Chen, D., 2019. Measurements and APSIM modelling of soil C and N dynamics. *Soil Res.* 58 (1), 41–61.
- Søgaard, H.T., Sommer, S.G., Hutchings, N.J., Huijsmans, J.F.M., Bussink, D.W., Nicholson, F., 2002. Ammonia volatilization from field-applied animal slurry - the ALFAM model. *Atmos. Environ.* 36 (20), 3309–3319.
- Signor, D., Cerri, C.E.P., 2013. Nitrous oxide emissions in agricultural soils: a review. *Pesquisa Agropecuaria Trop.* 43 (3), 322–338. <https://doi.org/10.1590/S1983-40632013000300014>.
- Stella, T., 2015. *Modelling Carbon and Nitrogen Dynamics in Paddy Rice system: Productivity and Greenhouse Gas emissions*. PhD Thesis. Università degli studi di Milano.
- Thompson, R.L., Lassaletta, L., Patra, P.K., et al., 2019. Acceleration of global N₂O emissions seen from two decades of atmospheric inversion. *Nat. Clim. Chang.* 9, 993–998. <https://doi.org/10.1038/s41558-019-0613-7>, 2019.
- Tian, H., Xu, R., Canadell, J.G., et al., 2020. A comprehensive quantification of global nitrous oxide sources and sinks. *Nature* 586, 248–256. <https://doi.org/10.1038/s41586-020-2780-0>.
- Vieten, B., Conen, F., Seth, B., Alewell, C., 2008. The fate of N₂O consumed in soils. *Biogeosciences* 5 (1), 129–132. <https://doi.org/10.5194/bg-5-129-2008>.
- Wang, Y., Guo, J., Vogt, R.D., Mulder, J., Wang, J., Zhang, X., 2018. Soil pH as the chief modifier for regional nitrous oxide emissions: new evidence and implications for global estimates and mitigation. *Glob. Chang. Biol.* 24, e617–e626. <https://doi.org/10.1111/gcb.13966>.
- Yin, X., Kersebaum, K.-C., Beaudoin, N., Constantin, J., Chen, F., Louarn, G., Manevski, K., Hoffmann, M., Kollas, C., Armas-Herrera, C.M., Baby, S., Bindi, M., Dibari, C., Ferchaud, F., Ferrise, R., de Cortazar-Atauri, I.G., Launay, M., Mary, B., Moriondo, M., Öztürk, I., Ruget, F., Sharif, B., Wachter-Ripoche, D., Olesen, J.E., 2020. Uncertainties in simulating N uptake, net N mineralization, soil mineral N and N leaching in European crop rotations using process-based models. *Field Crop. Res.* 255, 107863.
- Zhang, Q., Chen, M., Leng, Y., Wang, X., Fu, Y., Wang, D., Zhao, X., Gao, W., Li, N., Chen, X., Fan, C., Li, Q., 2023. Organic substitution stimulates ammonia oxidation-driven N₂O emissions by distinctively enriching keystone species of ammonia-oxidizing archaea and bacteria in tropical arable soils. *Sci. Tot. Environ.* 872, 162183 <https://doi.org/10.1016/j.scitotenv.2023.162183>.