Agricultural and Forest Meteorology 297 (2021) 108281

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Does agroecosystem model improvement increase simulation accuracy for agricultural N_2O emissions?

Yajie Zhang [a,b,*], Qiang Yu [a,b]

[a] State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Institute of Soil and Water Conservation, Northwest A&F University, Yangling, Shaanxi, 712100, China
[b] Institute of Soil and Water Conservation, Chinese Academy of Sciences and Ministry of Water Resources, Yangling, Shaanxi, 712100, China

ARTICLE INFO

Keywords:
Nitrous oxide emission
Agroecosystem model evolution
Coefficient of determination
Root mean square error
Meta-analysis

ABSTRACT

In recent decades, agroecosystem models have been developed to simulate agricultural nitrous oxide (N_2O) emissions. Coefficients of determination (R^2) and root mean square error (RMSE) are widely used as metrics to assess the explanatory power and simulation accuracy of models and to provide perspectives on model improvement as models evolve. This study aimed to determine whether the fitting accuracy of three agroecosystem models to simulate agricultural N_2O emissions improved with advancing versions of the models. We used several quality evaluation criteria to extract 94 and 97 reported R^2 and RMSE values, respectively, from 32 published articles related to the use of three of the most-used agroecosystem models in the research field of N_2O emissions [i.e., DNDC (DeNitrification-DeComposition), DayCent, and APSIM (Agricultural Production Systems sIMulator)]. Results showed that there was (1) no significant improvement in simulating N_2O emissions between DNDC9.3 and DNDC9.5; and (2) no significant difference between the simulation abilities of the original models and the user-defined revised models for these widely-used models. These findings may be mainly a result of the offsetting consequences of changes in publication bias and increased focus on complex agricultural issues. The study also found that the simulation accuracy of DNDC was better under conditions of higher annual mean temperature and soil bulk density and lower soil total nitrogen, mainly caused by the formulas and data used to build and validate the model. The study results suggest that the suitability of a model for simulating N_2O emissions depends on the climatic and soil conditions at the location of its application. Improving the simulation accuracy of agroecosystem models will require further targeted corrective and development actions in the future.

1. Introduction

Nitrous oxide (N_2O) is a greenhouse gas that has a long atmospheric lifetime (~121 yr) and a great global warming potential [100-yr global warming potential (GWP100) of 298; GWP20 of 268] (Aliyu et al., 2018; Xu-Ri et al., 2019). Agriculture is the largest anthropogenic source of N_2O emissions, accounting for 56–81% of the total gross anthropogenic emissions and 25–39% of the total global emissions (Davidson and Kanter, 2014; Aliyu et al., 2018). N_2O emissions from agriculture are determined by several factors, including climate and geographical conditions, soil properties, and agricultural management. To estimate the impacts of these factors on agricultural N_2O emissions, two approaches are usually taken: (1) empirical statistical models based on historical observation data; or (2) process-based agroecosystem models (Cannavo et al., 2008; Brilli et al., 2017). Compared with statistical regression models, process-based agroecosystem models generally simulate a suite of biogeochemical processes (e.g., ammonia volatilization, nitrification and denitrification, plant growth, organic matter decomposition, and fermentation), enabling computation of nitrogen transport and transformations in soil–plant–atmosphere ecosystem, and mechanistically representing the effects of different environmental variables and management measures on N_2O emissions (Brilli et al., 2017; Giltrap et al., 2020). In recent decades, these agroecosystem models have rapidly developed and been updated with parameter corrections, equation modifications, and sub-model development, thus improving the representation of agroecosystem dynamics and related research (e.g., Gilhespy et al., 2014; Holzworth et al., 2014). In most cases, the latest version of a model is directly applied in related research

* Corresponding author: State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Institute of Soil and Water Conservation, Northwest A&F University, Yangling, Shaanxi 712100, China.
E-mail addresses: zhangyajie1990@yeah.net (Y. Zhang), yuq@nwafu.edu.cn (Q. Yu).

https://doi.org/10.1016/j.agrformet.2020.108281
Received 20 June 2020; Received in revised form 2 December 2020; Accepted 6 December 2020
Available online 16 December 2020
0168-1923/

without comparing its simulation results with those of previous versions of the model. The explanatory power of models simulating N_2O emission variability is worthy of attention, and model version updates should be evaluated as to whether they improve simulation accuracy.

There are a number of metrics that have been developed and are widely used to assess the explanatory power of models. Two commonly used measures are the coefficient of determination (R^2) and the root mean square error (RMSE). The R^2 measures the amount of variation accounted for by the fitted model, while the RMSE measures the error between the fitted values with the observed data (Pham, 2019). An R^2 equal to 1 and a RMSE equal to 0 indicate perfect model fit. Although important discussions continue regarding the best methods for measuring the amount of variance explained by a model, the values of the R^2 and RMSE are often used to assess relative and absolute measures, respectively, of a model's fit of simulated results with observed measurements (Chai and Draxler, 2014; Low-Décarie et al., 2014). Low-Décarie et al. (2014) analyzed data from more than 18,000 articles and found that steadily falling R^2 values indicated a decrease in the explanatory power of ecology. According to reported R^2 values, Weisburd and Piquero (2008) evaluated how much explanatory power there was in criminological research, and found only a low level of variance was explained by the research, with no improvement over time. Yeluripati et al. (2015) evaluated the state of the DNDC (DeNitrification-DeComposition) crop model using a meta-analysis. They showed that the predictive power of accumulative annual N_2O emission fluctuations changed as DNDC version changed from 1.0 to 9.1, and that daily N_2O emissions were poorly modelled by DNDC. However, their study was limited by spatial and temporal scale heterogeneity, type of ecosystem modelled, and model complexity and development objectives in these modelling studies. We used R^2 and RMSE as measures of explanatory power and simulation accuracy of models in a meta-analysis study to answer the following questions: (1) what is the trend of R^2 and RMSE values for N_2O emissions simulation in relation to version updating of major agroecosystem models? (2) Is there a significant improvement in model performance using a user-defined revised version compared with the original version? (3) Do model revisions improve N_2O emission simulations across a range of crop types, geographical regions, and agricultural management practices? The results of this study may provide the impetus for continued work on revising and updating agroecosystem models.

2. Data and Methods

2.1. Data collection

A dataset was compiled based on published R^2 and RMSE values obtained from a literature survey of peer-reviewed publications in the Scopus abstract and citation database (https://www.scopus.com/home.uri) to identify English-language articles published before May 3, 2020. The search terms used were "agricultur*", "model*", and "N2O" or "nitrous oxide" in the subject areas of "environmental science", "agricultural and biological sciences", and "earth and planetary sciences". Based on the title, abstract, and keywords, studies were excluded that did not report agricultural N_2O emissions in studies of agroecosystem modeling. Three models appeared most frequently in this dataset and were regarded as the research subjects: DNDC, DayCent, and APSIM (Agricultural Production Systems sIMulator), with 142, 50, and 33 relevant papers, respectively, available for further processing (Giltrap et al., 2020). Different versions of DNDC were considered in this study, including models regionalized for different countries/regions (e.g., NZ-DNDC, UK-DNDC) and modified to suit specific crops or livestock farms (e.g., DNDC-Rice, DNDC-CSW, Manure-DNDC). A description of each model is provided in Table 1 and further details of the key differences and similarities with respect to the simulation of N_2O emissions can be found in Giltrap et al. (2020).

We followed the methods used in some previously published meta-analyses (e.g., Aliyu et al., 2018; Zhao et al., 2020), and we used the following criteria to select R^2 and RMSE values for use in this study: (1) the study was performed for identifiable field crops or forage grasses using at least one of the targeted models [model version(s) must be reported or could be inferred]; (2) N_2O emission treatments were randomized and replicated at least three times under field conditions and over at least two entire growing seasons for crops and two years for grasses; (3) N_2O emissions were measured hourly to biweekly using the static chamber method combined with the gas chromatography technique (Charteris et al., 2020); (4) at least one comparison of observed and simulated daily N_2O emissions (not cumulative or linearly interpolated emissions) was included; and (5) R^2 (or correlation coefficient), RMSE [or normalized RMSE, equal to 100 × RMSE / mean(observed)], sample size (n), and *p* value were provided or could be calculated.

2.2. Data pretreatment

Some previously published articles simultaneously reported R^2/RMSE values in more than one step of model application (e.g., model calibration, model validation, and model simulation). In contrast, only

Table 1
Details of the agroecosystem models compared in this study [modified from Cannavo et al. (2008) and Brilli et al. (2017)].

Model*	Factors influencing nitrous oxide production: Nitrification	Denitrification	Spatial scale	Time step	Website and software availability	Main references
DNDC	Nitrifier biomass, soil ammonium content, dissolved organic carbon content, soil temperature, soil water-filled pore space, and soil pH	Denitrifier biomass, total nitrogen as the sum of NO_3^-, NO_2^-, NO, and N_2O, dissolved organic carbon content, soil temperature, and soil pH	Field to regional	Daily (denitrification is calculated hourly following rainfall or irrigation events)	http://www.dndc.sr.unh.edu/; available online	(Li et al., 1992, Li et al., 1994); Li, 2000; DNDC, 2017
DayCent	Soil ammonium content, soil temperature, soil water-filled pore space, and soil pH	Soil nitrate content, soil respiration rate, and soil water-filled pore space	Field to global (ecosystems)	Daily	https://www2.nrel.colostate.edu/projects/daycent-home.html; available on request from authors (century@colostate.edu)	Del Grosso et al., 2000; Parton et al., 2001; Zhang et al., 2020
APSIM	Soil ammonium content, dissolved organic carbon content, soil temperature, soil water-filled pore space, and soil pH	Soil nitrate content, soil respiration rate, and soil water-filled pore space	Field to global	Mostly daily	https://www.apsim.info/; available online	Keating et al., 2003; Holzworth et al., 2014; APSIM, 2020a

* All three models included the following nitrogen processes: mineralization, leaching, volatilization, nitrification, denitrification, nitrogen uptake, and N_2 fixation.

one kind of R^2/RMSE value was recorded in our study using the following prioritization: R^2/RMSE from all data simulations > R^2/RMSE from validation > R^2/RMSE from calibration. If multiple publications contained data from the same study, only the R^2/RMSE values in the first study were retained. Models were categorized as "revised model" for comparison with the original model if they had been revised or integrated/coupled with other model(s) by model users (who were not the model developers) in some studies (R^2/RMSE values from both the original and revised models were recorded if available). If a study reported one value of R^2/RMSE using data from multiple experimental sites, the same or similar data for climate, soil, crop type, and management from close-distance sites were recorded. By following these criteria, 94 observations of R^2 from 25 publications and 97 observations of RMSE from 23 publications were extracted, covering 14 countries/regions. There were 16 publications that reported both R^2 and RMSE values. Because there was only one reported R^2/RMSE value for DNDC8.3P and DNDC9.2, these two models were not included in the dataset. Also, DNDCv.CAN (the Canadian version of DNDC that focused on cool-weather agriculture) was treated as a revised version of DNDC9.3 in this study. LandscapeDNDC unified DNDC9.3 and Forest-DNDC into a general soil biogeochemistry module in 2013, and was considered a separate version of DNDC in this analysis. This version was designed to simulate multi-ecosystems (i.e., forest, cropland, and grassland) and allowed the dynamic simulation of land use changes (Haas et al., 2013). For the DayCent model, there were several different versions (e.g., DayCent v4.5 2006, DayCent v4.5 2010, DayCent v4.5 2013, and DayCent v4.5 after 2013) resulted in different parameter settings and a few variations in the model structure (Sándor et al., 2018). All these different versions were treated as the "DayCent4.5" model in this study, because most of the related publications did not mentioned the specific versions used. Also, the development of the APSIM model went through different versions and generations. However, one-third of the related publications stated that APSIM7.5 was applied, but the others did not specify what model versions were used. Therefore, the designation of "APSIM7.x" was used in this study to represent the different versions of the APSIM model. The complete list of all changes and fixed defects to APSIM can be found in the release notes at https://www.apsim.info/download-apsim/downloads/. The details of these compiled peer-reviewed papers are listed in the Supplementary Information Table S1.

All data obtained from these papers [e.g., country/region, site location, climate conditions (annual mean air temperature and total precipitation), topsoil properties (pH, bulk density, clay content, organic carbon content, and total nitrogen content), crop types, agricultural management (total nitrogen application rate, irrigation rate, and tillage method), etc.] are available from the corresponding author upon reasonable request. When soil organic matter was reported during the dataset compilation process, it was converted to organic carbon content by dividing the value by 1.724 (Oldfield et al., 2019). Different studies reported topsoil properties at different depths. When studies reported topsoil properties at multiple depths, we averaged values of topsoil properties across depths to 20 cm (Oldfield et al., 2019). When studies reported topsoil properties to depths varying from 0 to 20 cm, we recorded values of topsoil properties directly without any conversion as suggested by Dr. Nils Borchard, Natural Resources Institute Finland (Luke) (https://www.researchgate.net/post/Is_there_an_acceptable_way_to_convert_soil_carbon_concentration_values_for_a_certain_depth_to_another_for_0-20cm_to_0-10cm).

2.3. Data analysis

The R^2 and RMSE values were assessed for data heterogeneity using the Levene's test, and for normality using the Shapiro-Wilk test. All data were log-transformed (if necessary) prior to analysis. Significant differences among the simulated R^2 and RMSE values of different model versions were tested by one-way ANOVA followed by the Tukey HSD mean separation test at $p < 0.05$. Simple linear and quadratic regression analyses were conducted to investigate the relationship of the R^2 and RMSE values to climate, topsoil, and agricultural management conditions. Data processes and statistical analyses were performed using R (version 3.3.1; Statistics Department of the University of Auckland; https://www.r-project.org/) and OriginPro8.5 (OriginLab Corporation; https://www.originlab.com/). The main R packages used were "stringr", "plyr", and "reshape2". Figures were created and regression analyses were conducted using OriginPro8.5.

3. Results

3.1. Comparisons between different agroecosystem models and model versions

Fig. 1 shows the R^2 and RMSE values between measured and simulated agricultural N_2O emissions for the six models/model versions. DNDC9.5 and DNDC9.3 had the largest mean R^2 values (0.422 and 0.414, respectively). LandscapeDNDC and DNDC9.3 had the smallest mean RMSE values (9.31 and 23.47 g N ha^{-1} d^{-1}, respectively). DNDC9.4 had significantly lower R^2 than DNDC9.5 and DNDC9.3 and larger RMSE than LandscapeDNDC and DNDC9.3 (p < 0.05). APSIM7.x and DayCent4.5 were created for different purposes than the DNDC model versions, and had lower mean R^2 values but smaller RMSE values for the simulation of N_2O emissions than some of the four DNDC versions. The R^2 and RMSE values for APSIM7.x and DayCent4.5 were not compared with each other, and will be discussed in the 4.4 Section.

Fig. 2 shows comparisons of the mean R^2 and RMSE values for the original and revised versions of DayCent4.5, LandscapeDNDC, and DNDC9.5. According to the R^2 values, for all three models, the simulations of N_2O emissions from the revised model did not significantly better fit the measured N_2O emissions than the original model did. Additionally, only the DayCent4.5 model had a numerically higher mean R^2 value for the revised model than for the original version (0.216 vs 0.150). Numerically lower mean RMSE values were observed for the revised DayCent4.5 and DNDC9.5 models than for the original versions, but none of the differences between versions of the same model were significant. Only LandscapeDNDC had a significantly higher mean RMSE value for the revised model than for the original version (33.33 vs 4.16 g N ha^{-1} d^{-1}).

3.2. Comparisons of the effect of different environmental and management factors on simulation accuracy of agroecosystem models

Figs. 3 and 4 show comparisons of R^2 and RMSE values for the four versions of DNDC for different environmental and management factors divided into two groups for each factor according to related threshold values. These thresholds were selected based on the representativeness of related environmental and management factor classifications (e.g., pH = 7 can classify soil into acid or alkaline soils). Only the differences of between-group (for the same model version) and intergroup (for different model versions) were compared. The results for the climate conditions (Figs. 3a and 4a, temperature, and 3b and 4b, precipitation) show that (1) most of the models simulated N_2O emissions better numerically under higher annual mean air temperature and under lower annual precipitation; (2) on the whole, DNDC9.3 performed better than the other versions of the model under the different climatic conditions; and (3) there were no significant differences for the between-group and intergroup comparisons of the R^2 values. For geographical conditions (Figs. 3c and 4c, latitude, and 3d and 4d, elevation), (1) most of the models simulated N_2O emissions with higher explanatory power but lower accuracy in mid-low latitudes; (2) DNDC9.3 showed relatively more stable performance than other versions of the model in the mid-low latitudes; but DNDC9.5 simulated N_2O emissions significantly better with higher R^2 and lower RMSE at mid-low latitudes than at mid-high latitudes; (3) DNDC9.3 had significantly higher mean R^2 and lower

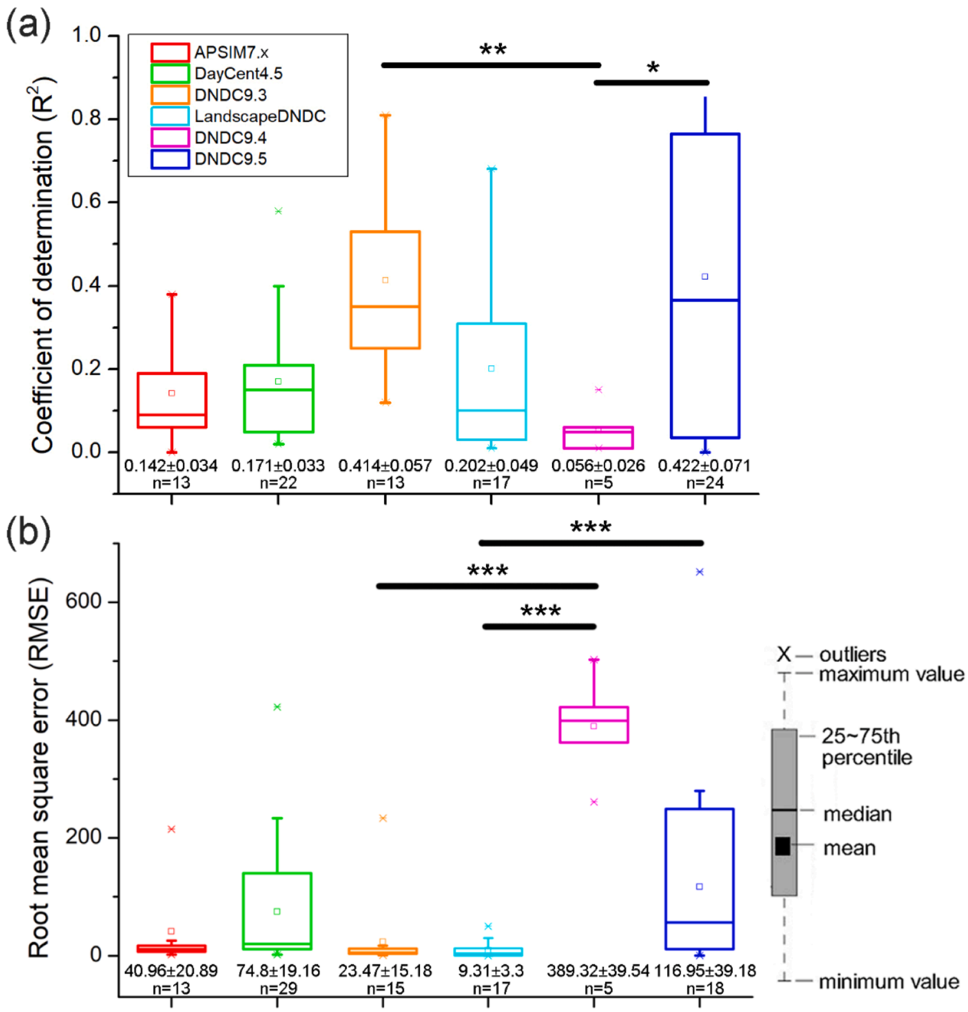

Fig. 1. Box plots of (a) coefficients of determination (R^2) and (b) root mean square error (RMSE; unit: g N ha^{-1} d^{-1}) between measured and simulated agricultural nitrous oxide emissions from publications reporting research using different agroecosystem models [DNDC (DeNitrification-DeComposition), DayCent, and APSIM (Agricultural Production Systems sIMulator)]. The lower and upper box boundaries indicate the 25th and 75th percentiles; the line inside the box indicates the median; the square inside the box indicates the mean; the lower and upper whiskers indicate the minimum and maximum values; the upper and lower X symbols beyond the ends of the whiskers indicate the outliers. Significant differences in R^2 and RMSE between model versions are indicated as *: $p < 0.05$ and **: $p < 0.01$. Values under the box plots are the mean ± standard error of the mean and the number of R^2 and RMSE values extracted from the publications.

mean RMSE than DNDC9.4 and DNDC9.5 at mid-high latitudes; and (4) DNDC9.5 simulated measured values better (higher R^2 and lower RMSE) at higher elevations than lower elevations, but the differences in mean R^2 and RMSE values were not significant. For agricultural management practices (Figs. 3e and 4e, nitrogen application rate, and 3f and 4f, irrigation amount), (1) DNDC9.3 showed numerically higher R^2 under higher nitrogen application rates than under lower rates, but the reverse was seen for DNDC9.5; (2) overall, DNDC9.5 had the numerically highest R^2 and the highest RMSE values under both nitrogen application conditions; and (3) explanatory power with DNDC9.5 was slightly, but non-significantly, greater with lower irrigation amounts. But simulation accuracy with DNDC9.5 showed opposite performance. For soil properties (Figs. 3g–k and 4g–k), (1) simulations with DNDC9.5 performed better (higher explanatory power and closer to measured values) when soil bulk density was greater than 1.2 g cm^{-3}, but the large differences in mean R^2 and RMSE values between the two bulk density categories were not statistically significant for this model version. In contrast, obvious (and significant in some cases) decreases in mean R^2 and increases in mean RMSE were observed for DNDC9.3 and DNDC9.5 when soil organic carbon was greater than 20 g kg^{-1} and total nitrogen content was greater than 2 g kg^{-1}; (2) DNDC9.3 and LandscapeDNDC performed significantly better when soil organic carbon and total nitrogen contents were at higher levels; and (3) both DNDC9.3 and DNDC9.5 performed better under higher soil pH conditions and higher clay content, but the increased performance in terms of R^2 was not statistically significant.

The effect of vegetative cover on explanatory power and simulation accuracy of N_2O emission varied from model to model (Figs. 5a and 5c). There were no significant differences between the R^2 and RMSE values in cropland and grassland from different model simulations. APSIM7.x performed better (but non-significantly) when used to simulate N_2O emissions over grassland than over cropland. Most of the DNDC versions showed numerically higher R^2 but higher RMSE when simulating N_2O emissions from cropland than from grassland. There was no significant difference between the mean R^2 and RMSE values obtained with either DayCent4.5 or DNDC9.5 due to tillage system (Figs. 5b and 5d). These two models explained N_2O emissions from no-till systems better than from conventional tillage, but had decreased simulation accuracy, as noted by the numerically greater R^2 but greater RMSE under no-till.

3.3. Relationships between R^2 (or RMSE) for measured and simulated N_2O emissions of widely-used agroecosystem models and different environmental and management factors

Obvious linear or polynomial relationships (defined as adjusted $R^2 \geq 0.2$, $p < 0.05$) were observed between the values of R^2 derived from studies using DNDC9.5 and annual mean air temperature, soil bulk density, soil organic carbon content, and soil total nitrogen content (Fig. 6). The fitted relationships for DNDC9.5 showed linear increasing trends with annual mean air temperature and soil bulk density in the range of 7 to 17°C and 0.8 to 1.6 g cm^{-3}, respectively (Figs. 6a and 6g). For soil organic carbon content, R^2 values from DNDC9.5 showed a falling-rising quadratic relationship with the minimum occurring at a soil organic carbon content of 40 g kg^{-1} (Fig. 6j). R^2 values from DNDC9.5 simulations of N_2O emissions decreased linearly with

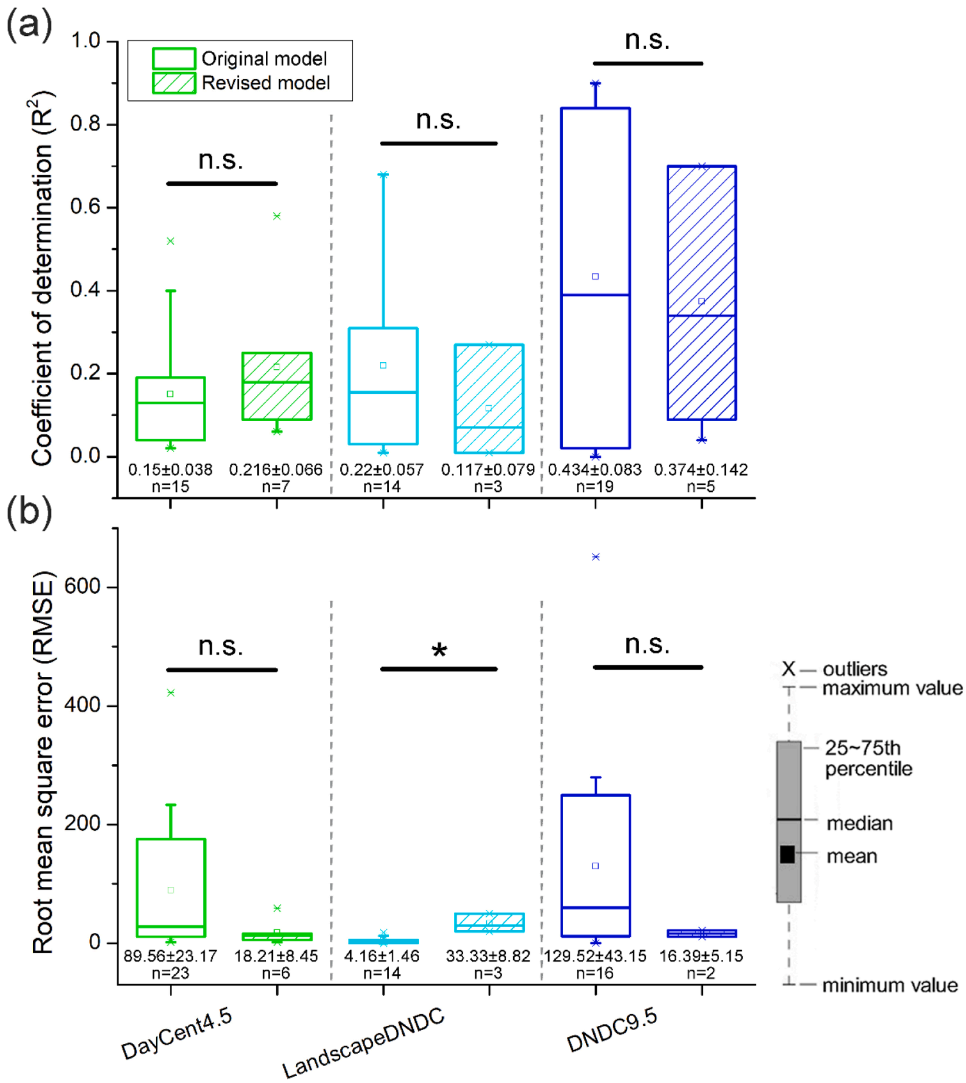

Fig. 2. Box plots of (a) coefficients of determination (R^2) and (b) root mean square error (RMSE; unit: g N ha^{-1} d^{-1}) between measured and simulated agricultural nitrous oxide emissions from publications reporting research using the original and revised versions of the DNDC (DeNitrification-DeComposition) and DayCent agroecosystem models. The lower and upper box boundaries indicate the 25th and 75th percentiles; the line inside the box indicates the median; the square inside the box indicates the mean; the lower and upper whiskers indicate the minimum and maximum values; the upper and lower X symbols beyond the ends of the whiskers indicate the outliers. n.s. denotes no significant difference at $p = 0.05$. Values under the box plots are the mean ± standard error of the mean and the number of R^2 and RMSE values extracted from the publications.

increasing soil total nitrogen content (Fig. 6k). There were, however, no clear trends in the relationships between the values of R^2 and any of the environmental and management factors for either the APSIM7.x or DayCent4.5 models.

Fig. 7 (RMSE) is presented similarly to Fig. 6 (R^2), and some main points should be noted: (1) APSIM7.x and DNDC9.5 showed higher simulation accuracy under higher annual mean air temperature (Fig. 7a). The DNDC9.5 results provided consistent information with Fig. 6a; (2) application of DayCent4.5 and DNDC9.5 in mid-low latitude areas may achieve better simulation results (Fig. 7c); (3) similar to the results presented in Fig. 6g, the higher the soil bulk density, the better the DNDC9.5 simulation (Fig. 7g); and (4) both DayCent4.5 and DNDC9.5 performed better at higher clay contents (Fig. 7i) or lower total nitrogen (Fig. 7k; consistent with the result shown in Fig. 6k for DNDC9.5) in agricultural soil.

4. Discussion

4.1. Simulation accuracy of N_2O emissions by agroecosystem models

Differences in physical and biogeochemical processes and input parameters used by these agroecosystem models likely contributed to the different results in different studies (Brilli et al., 2017). In DNDC, a dynamic "anaerobic balloon" function related to oxygen partial pressure was developed to allow nitrification and denitrification to occur simultaneously in aerobic or anaerobic microsites (Li et al., 2000). The simulation of free ammonium dynamics, nitrification, and nitrate leaching was subsequently further improved (Gilhespy et al., 2014). This increased the capacity to simulate urea hydrolysis for DNDC relative to the other models (Zimmermann et al., 2018). These functions improved the simulation of N_2O emission dynamics by DNDC. In APSIM, N_2O emission during nitrification was calculated following processes similar to those used by DNDC with a slightly different proportion of nitrified nitrogen (0.002 in APSIM vs 0.0024 in DNDC) and without consideration of a changing microbial population (Li, 2000; Vogeler et al., 2013; Holzworth et al., 2014; APSIM, 2020a). N_2O emission during denitrification was simulated by the model of Del Grosso et al. (2000), and was the same as the CENTURY biogeochemical model (DayCent was the daily time-step version of the CENTURY model) (Vogeler et al., 2013; Holzworth et al., 2014; APSIM, 2020a). The combination of these different formulas in APSIM resulted in a comparatively ideal simulation accuracy (mean RMSE was 40.96 g N ha^{-1} d^{-1}) but weak explanatory power of N_2O emissions (mean R^2 was 0.142). The same or opposite conclusions were drawn by some recent model comparison studies. For example, Gaillard et al. (2018) found DayCent simulations of N_2O emissions in agricultural fields of the USA were more significantly and strongly correlated with observed emissions than simulations produced with DNDC. Even though some versions of DNDC produced the best representation of measured N_2O emissions, research objectives and experimental conditions must still be considered when selecting a

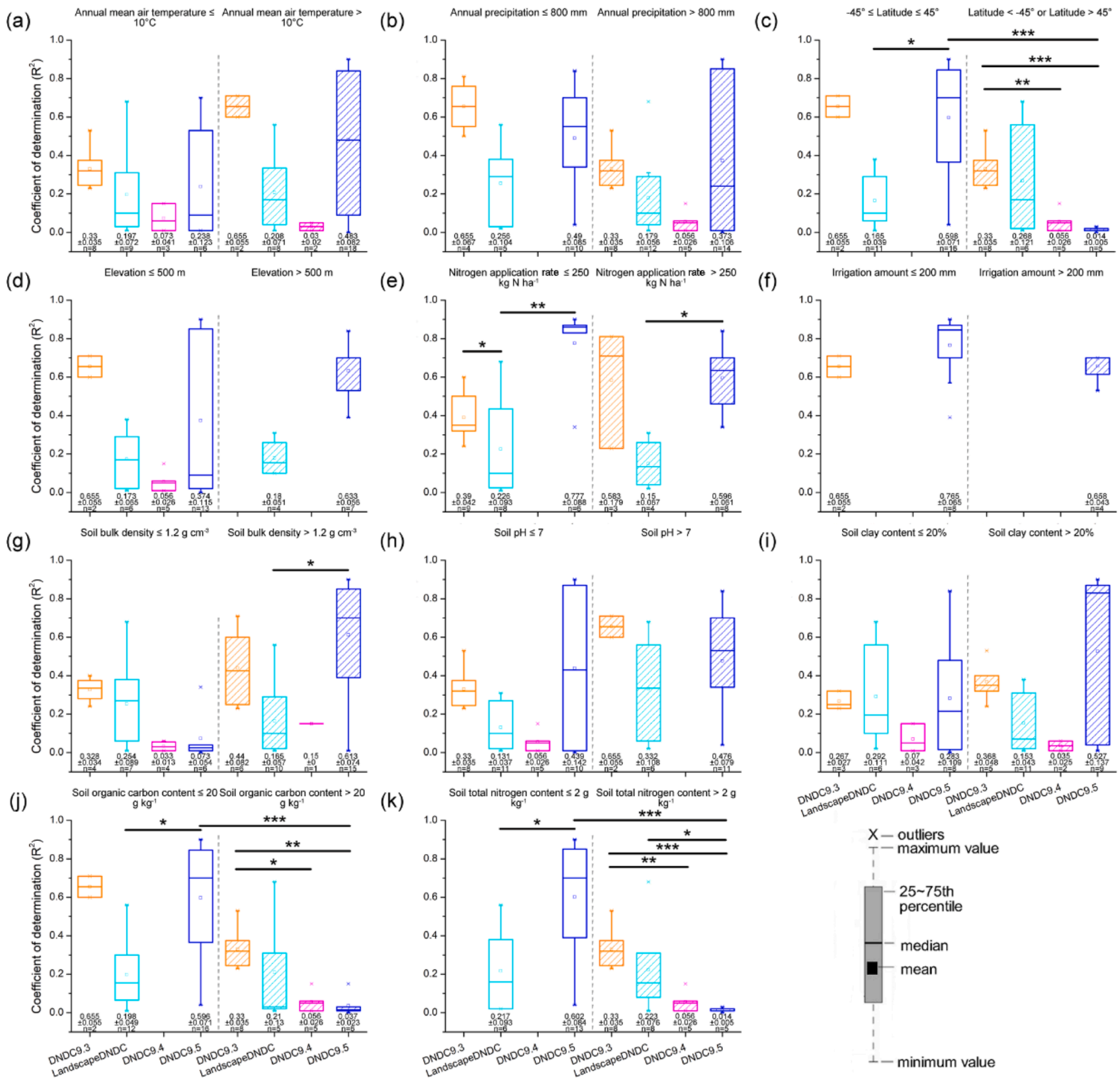

Fig. 3. Box plots of coefficients of determination (R^2) between measured and simulated agricultural nitrous oxide emissions from publications reporting research using the DNDC (DeNitrification-DeComposition) model for different (a–d) climate and geographical conditions, (e–f) agricultural management practices, and (g–k) soil properties. All of these environmental factors were divided into two groups according to related threshold values. The lower and upper box boundaries indicate the 25th and 75th percentiles; the line inside the box indicates the median; the square inside the box indicates the mean; the lower and upper whiskers indicate the minimum and maximum values; the upper and lower X symbols beyond the ends of the whiskers indicate the outliers. Only the differences of between-group (for the same model version) and intergroup are compared. Significant differences in R^2 between model versions are indicated as *: $p < 0.05$, **: $p < 0.01$, and ***: $p < 0.001$; values under the box plots are the mean ± standard error of the mean and the number of R^2 values extracted from the publications.

model.

Improved representation of soil water evaporation was incorporated into DNDC during its development (DNDC9.3) in 2010. The soil ammonia volatilization calculation process developed in Manure-DNDC was also used in DNDC9.4 in 2012 (Gilhespy et al., 2014). Simulating procedures of crop growth, hydrological cycle, and greenhouse gas production were further optimized for DNDC9.5 in 2014 (Gilhespy et al., 2014; Zhang and Niu, 2016). In contrast to the results seen for DNDC9.3 and DNDC9.5, DNDC9.4 failed to adequately simulate agricultural N_2O emissions. Most of the R^2 and RMSE observations for DNDC9.4 were derived from Zimmermann et al. (2018), who simulated N_2O fluxes from different grass and arable sites by DayCent, DNDC9.4, and DNDC9.5 in the Northern Ireland and Republic of Ireland. The models used in Zimmermann et al. (2018) showed overall worse performance than previous similar studies carried out in the UK and Ireland. However, because of the limited availability of related research, it is difficult to thoroughly explain why these differences in simulation results between related models occurred. Some potential explanations for why there has been no significant improvement in R^2 and RMSE values as new versions of the model have been created are: (1) agroecosystem models may have started to be used to answer more difficult questions regarding agricultural biogeochemical cycles and sustainable development that are less amenable to explanation; (2) an increasing number of model users who are not well-trained in agroecosystem modelling may result in poor application of the model because of not adequately understanding the model's theories, structures, data requirements, implementation, and limitations (Weisburd and Piquero, 2008); (3) the improvements made in new versions of a model also increased model complexity and possibly output sensitivity to inputs and model parameters; and (4) there may have been a shift in publication bias that acted directly on R^2 and RMSE values. For example, more studies that report low R^2 and high RMSE values may have been published due to the growing recognition of the importance of publishing "negative results" (Low-Décarie et al., 2014). Moreover, more researchers may report real research results under the

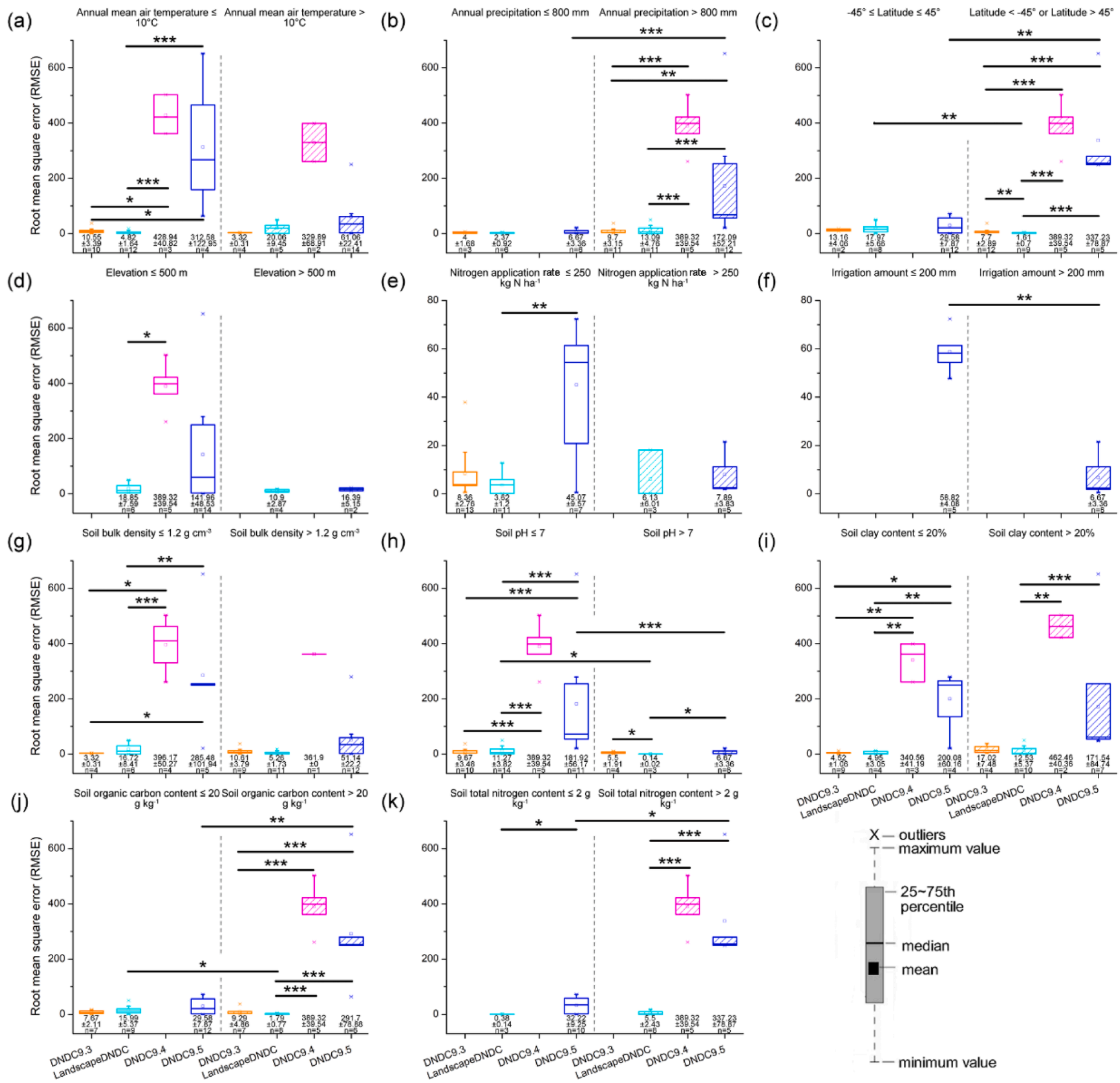

Fig. 4. Box plots of root mean square error (RMSE; unit: g N ha^{-1} d^{-1}) between measured and simulated agricultural nitrous oxide emissions from publications reporting research using the DNDC (DeNitrification-DeComposition) model for different (a–d) climate and geographical conditions, (e–f) agricultural management practices, and (g–k) soil properties. All of these environmental factors were divided into two groups according to related threshold values. The lower and upper box boundaries indicate the 25th and 75th percentiles; the line inside the box indicates the median; the square inside the box indicates the mean; the lower and upper whiskers indicate the minimum and maximum values; the upper and lower X symbols beyond the ends of the whiskers indicate the outliers. Only the differences of between-group (for the same model version) and intergroup are compared. Significant differences in RMSE between model versions are indicated as *: $p < 0.05$, **: $p < 0.01$, and ***: $p < 0.001$; values under the box plots are the mean ± standard error of the mean and the number of RMSE values extracted from the publications.

constraints of highlighted academic ethics and integrity in recent years.

Some researchers have tried to apply user-defined revised models that changed some internal formulas/parameters or were integrated/coupled with other model(s) to improve model performance under their specific conditions. For example, Yu et al. (2017) applied a revised DNDC9.5 model that improved simulations of N_2O emissions and crop growth when using plastic mulching in cotton fields in an arid region of China. The results shown in Fig. 2 indicated that for one of the three models, the revised model performed better than the original model, showing higher R^2 and lower RMSE. However, the differences were not significant ($p > 0.05$). Please note that the data used in this analysis were not the pairwise data because some publications did not report the fitting effect of model(s) before being revised to a new version, thereby causing selective reporting bias. A reason for why we did not observe consistent significant improvement in simulation accuracy with revised versions of the models compared with original versions may be partly due to the offsetting consequences of attempting to solve complex agricultural questions with models (leading to decreasing R^2 and increasing RMSE for revised models) and publication bias resulting from authors not publishing their results if a revised version of the model did not achieve a larger R^2 or lower RMSE value than the original model (leading to increasing R^2 and decreasing RMSE for revised models) (Thornton and Lee, 2000; Low-Décarie et al., 2014). However, studies using revised models have encouraged the development and evolution of models. A great deal of additional work will be needed to carefully consider the methods of model modification needed to achieve satisfactory results.

4.2. Relationships between environmental and management factors and values of R^2 and RMSE simulated by agroecosystem models

Although the formulas used in these agroecosystem models have been presented in some publications (e.g., DNDC, 2017; APSIM, 2020a; Zhang et al., 2020), it was still difficult to explain the relationships

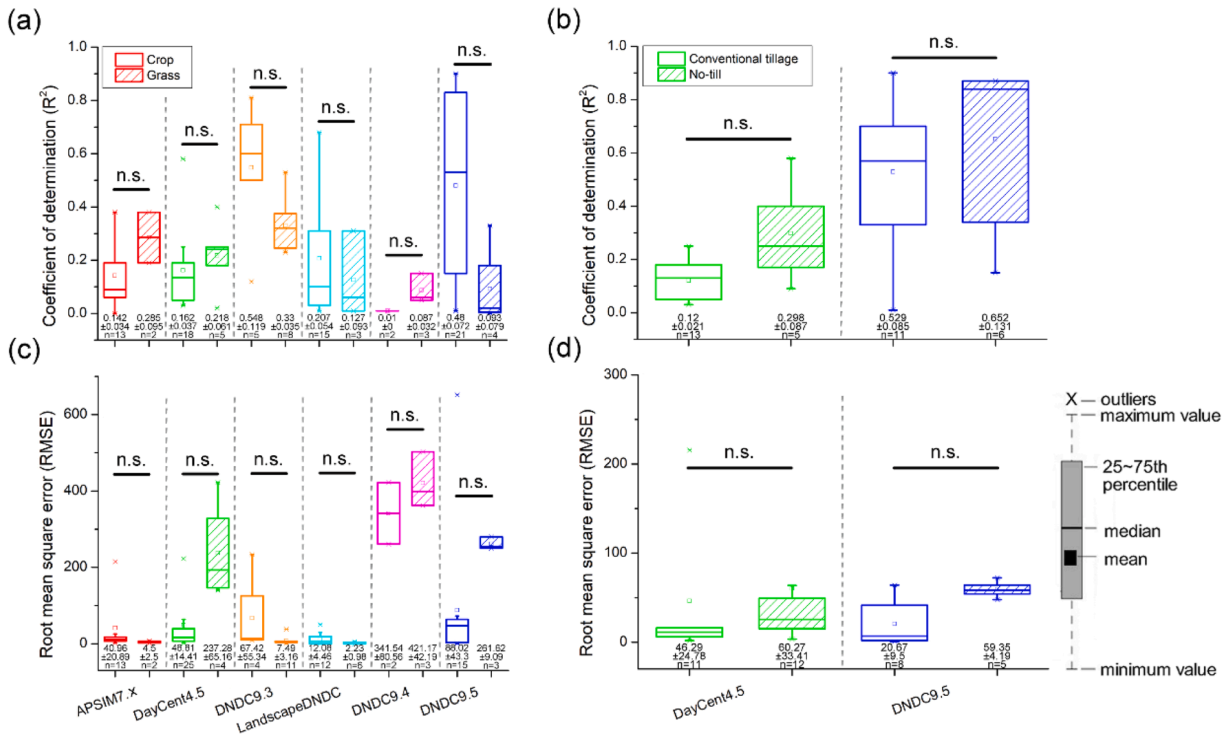

Fig. 5. Box plots of (a and b) coefficients of determination (R^2) and (c and d) root mean square error (RMSE; unit: g N ha^{-1} d^{-1}) between measured and simulated agricultural nitrous oxide emissions from publications reporting research using different agroecosystem models [DNDC (DeNitrification-DeComposition), DayCent, and APSIM (Agricultural Production Systems sIMulator)] for (a and c) cropland and grassland sites and for (b and d) conventional tillage and no-till. The lower and upper box boundaries indicate the 25th and 75th percentiles; the line inside the box indicates the median; the square inside the box indicates the mean; the lower and upper whiskers indicate the minimum and maximum values; the upper and lower X symbols beyond the ends of the whiskers indicate the outliers. n.s. denotes no significant difference ($p = 0.05$) in R^2 and RMSE values due to vegetative cover category or due to tillage category; Values under the box plots are the mean ± standard error of the mean and the number of R^2 and RMSE values extracted from the publications.

between the environmental and management factors and simulated N_2O emissions, and consequently the simulation accuracy of models. This was because (1) environmental and management factors affect N_2O emissions in nitrification and denitrification processes in diverse ways, (2) observational data on soil N_2O emissions contain significant temporal and spatial uncertainty, and (3) the processes of data collection and model operation may have uncertainties caused by the researchers (Myrgiotis et al., 2018). Some evidence regarding the relationship between environmental and management factors and simulated N_2O emissions have been provided by some studies using sensitivity analysis methods (Giltrap et al., 2010). Myrgiotis et al. (2018) examined the parameter sensitivity of N_2O simulated by LandscapeDNDC and identified the soil microbial dynamics, pH, soil porosity, and dissolved organic carbon content as the key sources of output sensitivity. Necpálová et al. (2015) examined the sensitivity of N_2O simulated by DayCent through inverse modeling using the PEST parameter estimation software and found that soil temperature, carbon/nitrogen ratio, nitrogen availability, and potential evapotranspiration noticeably affected model outputs. A comparison study showed that temperature and water content had larger nitrification effects in APSIM than in DNDC, while temperature and organic carbon content produced larger denitrication responses in DNDC than in APSIM (Vogeler et al., 2013). Vogeler et al. (2013) also indicated that DNDC simulated linearly increasing N_2O emission rates with increasing nitrogen load, while lower emission rates were simulated by APSIM. Also, increasing rainfall intensity increased N_2O emissions simulated by DNDC, but decreased emissions simulated by APSIM. These parameters made up the majority of the model's parameters, thereby affecting the size and state of the nitrogen-based soil substrate (Myrgiotis et al., 2018). There were some obvious signs that changes in the main parameters mentioned above may influence the R^2 and RMSE values simulated by these models (see Figs. 3, 4, 6, and 7).

Moreover, the relationships between the environmental and management factors and the simulation accuracy of the models may be considerably influenced by model structure, calibration, and validation. The data used in this study came from a limited range of soil/climate conditions and production systems. For example, Del Grosso et al. (2000) provided a general model that was used for simulating denitrification in APSIM and DayCent. This general model was developed based on collected soil core samples and repacked soil data from Weier et al. (1993) from northern Colorado. Thus APSIM and DayCent can effectively simulate the denitrification process and corresponding N_2O emissions from agroecosystems of the semi-arid Great Plains of the USA and other locations having similar environmental and agricultural conditions. The DNDC model was originally developed to simulate N_2O emissions from cropped soils in the USA. The calculation formulas for N_2O emissions coming from nitrification and denitrification were mainly sourced from many publications conducted in the USA, China, and some European countries (Li et al., 1992; Li, 2000). Therefore, most of the DNDC versions can generally produce good simulations at sites having climate (including irrigation events that are always treated as precipitation) and soil characteristics falling within the ranges of the original data used to develop DNDC.

In this study, we found that most of the DNDC versions were good at simulating N_2O emissions in croplands, while APSIM7.x and DayCent4.5 performed better for grasslands. Fuchs et al. (2020) conducted a multi-model evaluation in an intensively managed grassland in Switzerland, and indicated that DayCent and APSIM performed well for simulating annual and daily N_2O fluxes, respectively. Abdalla et al. (2010) indicated that DayCent simulated daily N_2O fluxes more consistently than DNDC did for pasture in Ireland. This was because the plant growth sub-model used in DayCent simulates the growth of various grasses and trees (Del Grosso et al., 2001). A pasture growth model

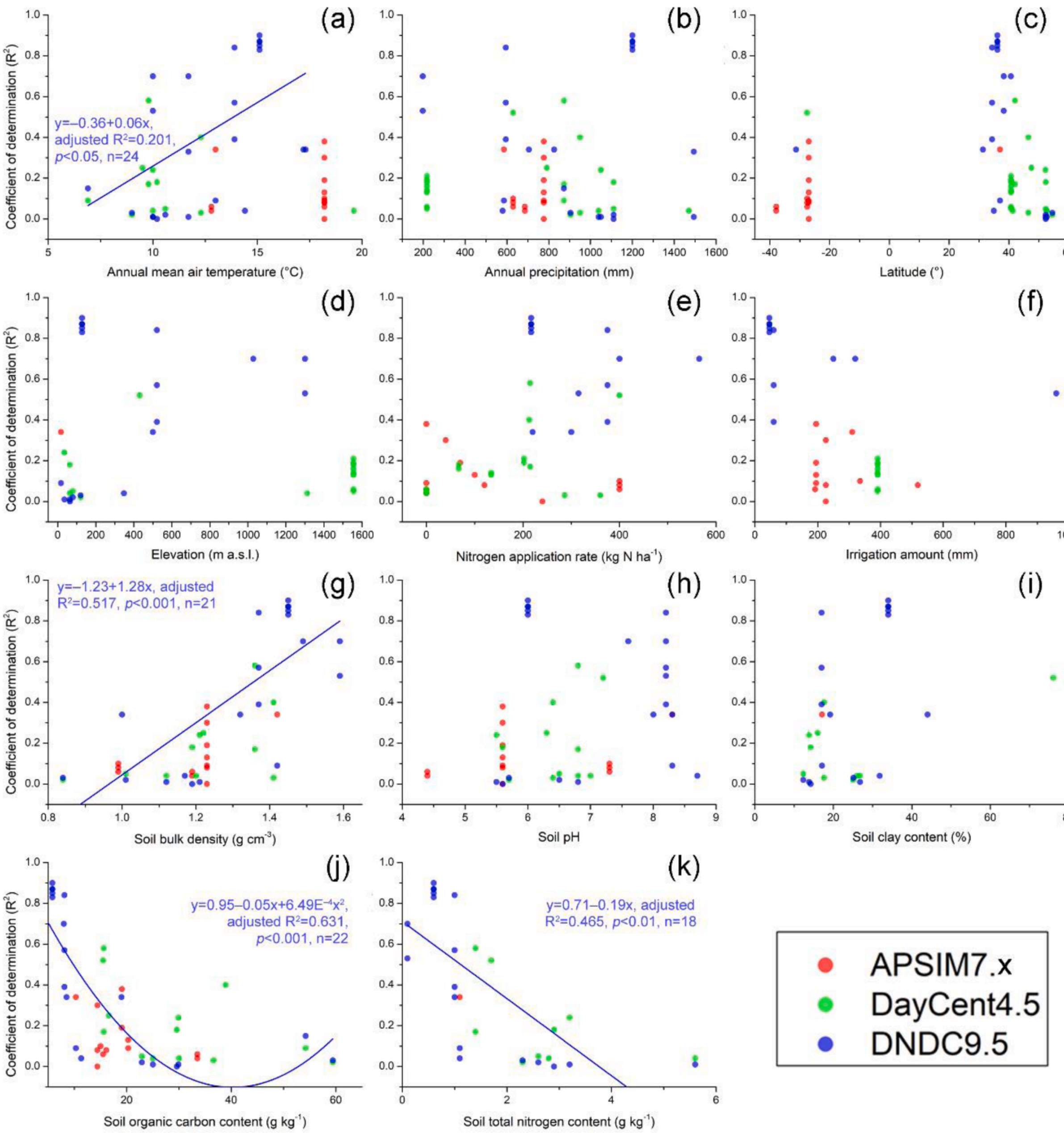

Fig. 6. Relationships between coefficients of determination (R^2) from publications reporting measured and simulated agricultural nitrous oxide emissions for (a–d) climate and geographical conditions, (e–f) agricultural management practices, or (g–k) soil properties using widely-used agroecosystem models [DNDC (DeNitrification-DeComposition), DayCent, and APSIM (Agricultural Production Systems sIMulator)]. The best fitting linear or quadratic regressions are presented when significant (adjusted $R^2 \geq 0.2$, $p < 0.05$), that only occurred for DNDC9.5.

(AgPasture) was developed and integrated into APSIM with a set of management tools for modelling pasture management such as grazing, cutting, and renewal (Holzworth et al., 2014; APSIM, 2020b). Some specific DNDC model versions (e.g., NZ-DNDC and Manure-DNDC) may be better choices for simulation of N_2O emissions from grasslands or pasture systems (Gilhespy et al., 2014). DNDC has been applied in many land use types, but the majority of the applications have been in croplands (Yeluripati et al., 2015).

4.3. What needs to be done in agroecosystem model development?

Climate change is causing increases in temperature, changes in precipitation, imbalances in hydrology and water resources, and increased frequency of extreme weather events (Trenberth, 2011; Cook et al., 2018). Increasing global agricultural intensification has increased the attention given to developing methods to conserve water resources and control nitrogen applications in order to ensure future food security. For now, the agroecosystem models discussed in this paper can achieve good simulations of N_2O emissions for conditions of higher temperature and lower nitrogen fertilizer application rates. However, only 15% of studies dealing with DNDC have focused on quantification of the impact of temperature (and precipitation) changes on model results (Yeluripati et al., 2015). Additionally, the influence of extreme weather events (e.g., drought, heatwave) on simulated N_2O emissions remains unclear. Under the influence of climate change, the usability or applicability of the formulas and parameters used in these models becomes more uncertain. For example, the effects of soil pH and temperature on denitrification are modeled in DNDC based on observations that were reported in a number

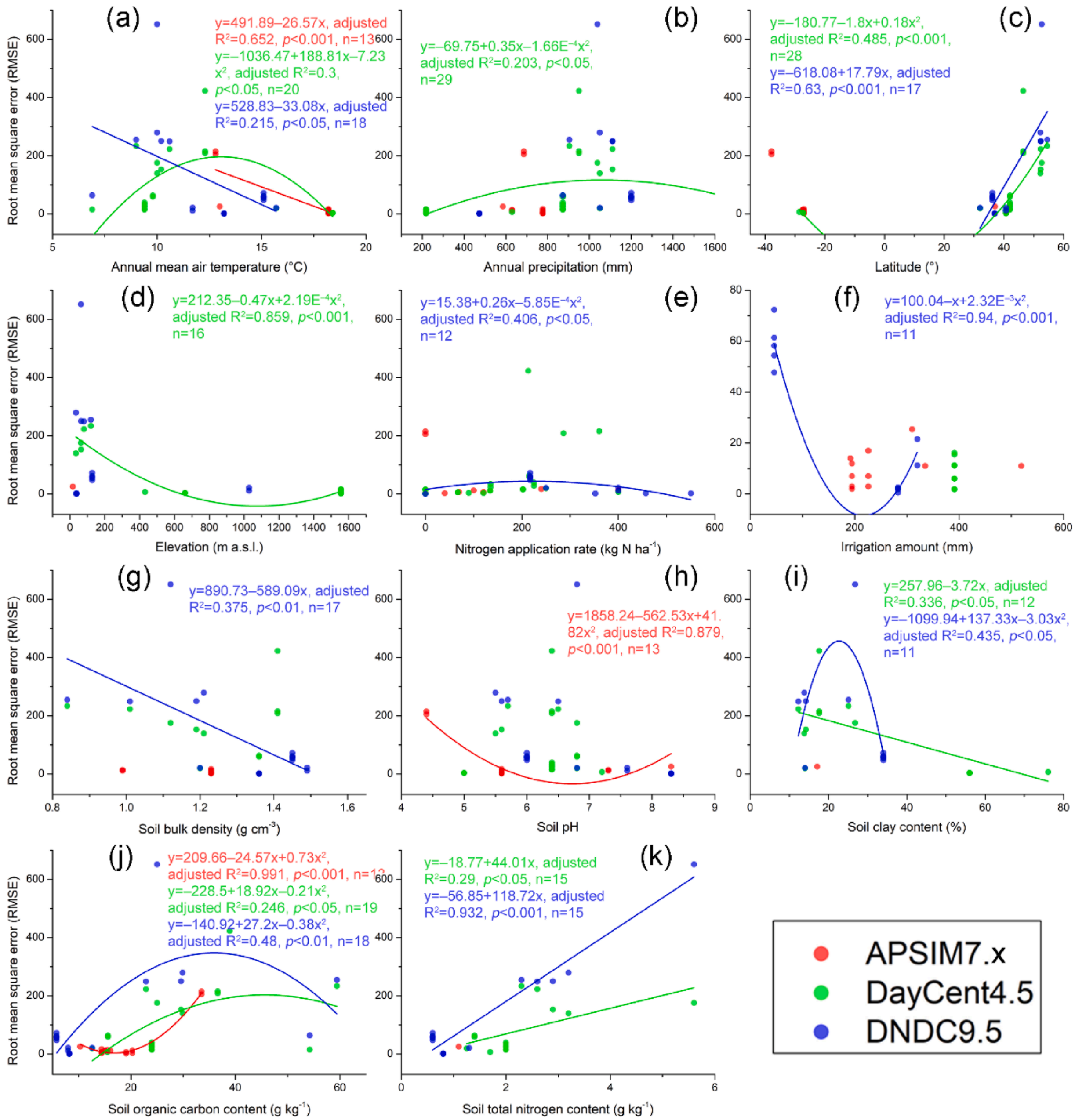

Fig. 7. Relationships between root mean square error (RMSE; unit: g N ha^{-1} d^{-1}) from publications reporting measured and simulated agricultural nitrous oxide emissions for (a–d) climate and geographical conditions, (e–f) agricultural management practices, or (g–k) soil properties using widely-used agroecosystem models [DNDC (DeNitrification-DeComposition), DayCent, and APSIM (Agricultural Production Systems sIMulator)]. The best fitting linear or quadratic regressions are presented when significant (adjusted $R^2 \geq 0.2$, $p < 0.05$).

of studies published from 1954 to 1988 (DNDC, 2017). Furthermore, more studies should be conducted at low-latitude agricultural sites to fill the research gap indicated in Fig. 6c (Zhang and Yu, 2020a). Overall, mean R^2 values were low (< 0.5) for all of the agroecosystem models investigated in this study (Fig. 1a). Coupling agroecosystem models and machine learning algorithms can help improve prediction and accelerate progress toward developing dynamic decision support tools for agricultural management (e.g., Shahhosseini et al., 2019). Great opportunities remain for future model development to further improve understanding of agricultural nitrogen cycle responses to changing climate and socioecological influences caused by nitrogen pollution in order to achieve realistic solutions to the problems faced by agricultural production systems in connection with climate change (Robertson et al., 2013).

4.4. Limitations of this study

This study had several potential limitations: (1) the dataset from Scopus did not include all journals available worldwide. Additionally, only journal articles were included in this study, without consideration of reviews, books, dissertations, and conference proceedings, etc., that also may have reported on model use and could have provided additional values of R^2 and RMSE (Zhang and Yu, 2020b). This may have limited the sample size of this study. (2) The ANOVA will technically work when having one value more than the number of groups for the total sample size. However, depending on the variability of the data, there still were concerns about the statistical power in ANOVA in many cases with very small sample sizes. (3) According to the publication year, mentioned model functions, and publishing records by the same authors, model versions were inferred when there were no related detailed descriptions in some publications. This may have caused some mis-estimating in this study. (4) Some included studies did not report site elevation and soil properties in detail. Using the study's coordinates, these data were extracted from some datasets (e.g., ISRIC SoilGrids) or from other papers that were conducted at the same trial site in close years to fill in the missing values of elevation and soil properties, but doing so may have affected the results presented in Figs. 3, 4, 6, and 7 (Oldfield et al., 2019). Moreover, some high-efficiency agricultural management technologies and methods (e.g., biochar, plastic film mulching) have been widely applied in actual agricultural production systems, but have not been included in some of the models' structure and calculations. Some related studies were retained in this analysis but may have further affected the results of this study. (5) R^2 and RMSE were used in this study to compare different models by assessing the ''goodness of fit'' of simulated N_2O emissions to measured values. However, these metrics only assessed how well a model predicted the emissions on a given day (Giltrap et al., 2010). Model predictions sometimes either led or lagged behind the observations by a few days, and sometimes made the model appear to perform poorly in terms of R^2 or RMSE. However, the model would still be reliable and useful for simulating accumulative long-term emissions, conducting scenario analysis to identify the main model drivers and optimal management measures, and estimating the future emissions under climate change scenarios (Giltrap et al., 2010). The R^2 and RMSE between measured and simulated values are not the sole criteria of success in model development and application (Low-Décarie et al., 2014). R^2 can be replaced by the adjusted R^2 to take into account the number of explanatory variables in a model relative to the number of data points (Pham, 2019). In addition, R^2 and RMSE are not irrelevant as $R^2 = 1 - RMSE^2$ / Variance, and it can be expected that R^2 will decrease as RMSE increases. Moreover, please note that the R^2 and RMSE values for different versions of DNDC can be compared due to their close release dates and numbers of predictor variables. But the R^2 of different agroecosystem models cannot be compared because different numbers of predictor variables may affect the values of R^2 for these different models. Though the RMSE of different agroecosystem models can be compared because their errors are measured in the same units, this study does not provide these RMSE comparisons because comprehensive conclusions can only be reached based on the comparisons of both R^2 and RMSE for these different models. (6) Because of the difficulty in accessing information on the numbers of parameters and formulas used in specific models (especially for the nonlinear equations), and information relative to conceptual process differences in models, it is difficult to thoroughly understand and illustrate why the performance of different models or model versions is improved or reduced. Investigating how the changing complexity of agroecosystem models affects their explanatory power and simulation accuracy over time is worthy of further investigation and study (Giltrap et al., 2020).

5. Conclusions

This analysis of published research provided useful insights into the current state and progress of simulations of N_2O emissions by agroecosystem models, but more work is warranted: (1) for APSIM and DayCent, there is still a great deal of room for improving simulation ability of N_2O emission dynamics from croplands; (2) for new versions of DNDC, the simulation accuracy of N_2O emission should be further improved; (3) the formulas and parameters used in these model development are somewhat out-of-date and need to be continuously evaluated under future climate change conditions; and (4) a combination of metrics, including but not limited to R^2 and RMSE, are encouraged to report for assessing model performance (Chai and Draxler, 2014). Model developers and researchers need to give much more attention to what the models can or cannot predict, what is not explained, and what implications result from these unexplained variations in modeling of agricultural N_2O emissions. Future research should expand beyond the scope of this work to provide scientific values for the development of agroecosystem 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.

Acknowledgements

This work uses Scopus data provided by Elsevier B.V. (https://www.scopus.com/search/form.uri?display=basic). The authors are very grateful to the editor and two anonymous reviewers for their valuable comments that have significantly improved the paper. The authors thank the large number of scientists who have developed models investigated in this study and the authors of previous publications that provided contributions to the dataset used in this study. The authors thank Dr. David C. Nielsen at AEREA Inc. (www.aereainc.com), and Mr. Shouhua Xu at the Institute of Soil and Water Conservation, Northwest A&F University for providing suggestions on the manuscript. This study was funded by the National Natural Science Foundation of China (No. 41961124006, 41730645), the International Partnership Program of the Chinese Academy of Sciences (No. 161461KYSB20170013), and the "111 Project" (No. B12007) of China.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.agrformet.2020.108281.

References

Abdalla, M., Jones, M., Yeluripati, J., Smith, P., Burke, J., Williams, M., 2010. Testing DayCent and DNDC model simulations of N_2O fluxes and assessing the impacts of climate change on the gas flux and biomass production from a humid pasture. Atmospheric Environment 44, 2961–2970.

Aliyu, G., Sanz-Cobena, A., Müller, C., Zaman, M., Luo, J., Liu, D., Yuan, J., Chen, Z., Niu, Y., Arowolo, A., Ding, W., 2018. A meta-analysis of soil background N_2O emissions from croplands in China shows variation among climatic zones. Agriculture, Ecosystems and Environment 267, 63–73.

APSIM, 2020a. APSIM 7.10: Soil modules documentation: SoilN, The APSIM Initiative, https://www.apsim.info/documentation/model-documentation/soil-modules-documentation/soiln/ (accessed June 2020).

APSIM, 2020b. APSIM 7.10: Crop module documentation: AgPasture, The APSIM Initiative, https://www.apsim.info/documentation/model-documentation/crop-module-documentation/agpasture/ (accessed June 2020).

Brilli, L., Bechini, L., Bindi, M., Carozzi, M., Cavalli, D., Conant, R., Dorich, C., Doro, L., Ehrhardt, F., Farina, R., Ferrise, R., Fitton, N., Francaviglia, R., Grace, P., Iocola, I., Klumpp, K., Léonard, J., Martin, R., Massad, R., Recous, S., Seddaiu, G., Sharp, J., Smith, P., Smith, W., Soussana, J.-F., Bellocchi, G., 2017. Review and analysis of strengths and weaknesses of agro-ecosystem models for simulating C and N fluxes. Science of the Total Environment 598, 445–470.

Cannavo, P., Recous, S., Parnaudeau, V., Reau, R., 2008. Modeling N dynamics to assess environmental impacts of cropped soils. Advances in Agronomy 97, 131–174.

Chai, T., Draxler, R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. Geoscientific Model Development 7, 1247–1250.

Charteris, A., Chadwick, D., Thorman, R., Vallejo, A., de Klein, C., Rochette, P., Cardenas, L., 2020. Global Research Alliance N_2O chamber methodology guidelines: Recommendations for deployment and accounting for sources of variability. Journal of Environmental Quality. https://doi.org/10.1002/jeq2.20126.

Cook, B., Mankin, J., Anchukaitis, K., 2018. Climate change and drought: From past to future. Current Climate Change Reports 4, 164–179.

Davidson, E., Kanter, D., 2014. Inventories and scenarios of nitrous oxide emissions. Environmental Research Letters 9, 105012. https://doi.org/10.1088/1748-9326/9/10/105012.

Del Grosso, S., Parton, W., Mosier, A., Hartman, M., Keough, C., Peterson, G., Ojima, D., Schimel, D., 2001. Simulated effects of land use, soil texture, and precipitation on N gas emissions using DAYCENT. In: Follett, R.F., Hatfield, R.F., J.L. (Eds.), Nitrogen in the Environment: Sources, Problems, and Management. Elsevier Science Publishers, The Netherlands, pp. 413–431.

Del Grosso, S., Parton, W., Mosier, A., Ojima, D., Kulmala, A., Phongpan, S., 2000. General model for N_2O and N_2 gas emissions from soils due to denitrification. Global Biogeochemical Cycles 14, 1045–1060.

DNDC, 2017. DNDC (version 9.5): Scientific basis and processes, Institute for the Study of Earth, Oceans, and Space of the University of New Hampshire, http://www.dndc.sr.unh.edu/papers/DNDC_Scientific_Basis_and_Processes.pdf (accessed June 2017).
Fuchs, K., Merbold, L., Buchmann, N., Bretscher, D., Brilli, L., Fitton, N., Topp, C., Klumpp, K., Lieffering, M., Martin, R., Newton, P., Rees, R., Rolinski, S., Smith, P., Snow, V., 2020. Multimodel evaluation of nitrous oxide emissions from an intensively managed grassland. Journal of Geophysical Research: Biogeosciences 125, e2019JG005261. https://doi.org/10.1029/2019JG005261.
Gaillard, R., Jones, C., Ingraham, P., Collier, S., Izaurralde, R., Jokela, W., Osterholz, W., Salas, W., Vadas, P., Ruark, M., 2018. Underestimation of N_2O emissions in a comparison of the DayCent, DNDC, and EPIC models. Ecological Applications 28, 694–708.
Gilhespy, S., Anthony, S., Cardenas, L., Chadwick, D., del Prado, A., Li, C., Misselbrook, T., Rees, R., Salas, W., Sanz-Cobena, A., Smith, P., Tilston, E., Topp, C., Vetter, S., Yeluripati, J., 2014. First 20 years of DNDC (DeNitrification DeComposition): Model evolution. Ecological Modelling 292, 51–62.
Giltrap, D., Li, C., Saggar, S., 2010. DNDC: A process-based model of greenhouse gas fluxes from agricultural soils. Agriculture, Ecosystems and Environment 136, 292–300.
Giltrap, D., Yeluripati, J., Smith, P., Fitton, N., Smith, W., Grant, B., Dorich, C., Deng, J., Topp, C., Abdalla, M., Liáng, L., Snow, V., 2020. Global Research Alliance N_2O chamber methodology guidelines: Summary of modeling approaches. Journal of Environmental Quality. https://doi.org/10.1002/jeq2.20119.
Haas, E., Klatt, S., Froehlich, A., Kraft, P., Werner, C., Kiese, R., Grote, R., Breuer, L., Butterbach-Bahl, K., 2013. LandscapeDNDC: a process model for simulation of biosphere–atmosphere–hydrosphere exchange processes at site and regional scale. Landscape Ecology 28, 615–636.
Holzworth, D., Huth, N., deVoil, P., Zurcher, E., Herrmann, N., McLean, G., Chenu, K., van Oosterom, E., Snow, V., Murphy, C., Moore, A., Brown, H., Whish, J., Verrall, S., Fainges, J., Bell, L., Peake, A., Poulton, P., Hochman, Z., Thorburn, P., Gaydon, D., Dalgliesh, N., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp, J., Cichota, R., Vogeler, I., Li, F., Wang, E., Hammer, G., Robertson, M., Dimes, J., Whitbread, A., Hunt, J., van Rees, H., McClelland, T., Carberry, P., Hargreaves, J., MacLeod, N., McDonald, C., Harsdorf, J., Wedgwood, S., Keating, B., 2014. APSIM – Evolution towards a new generation of agricultural systems simulation. Environmental Modelling & Software 62, 327–350.
Keating, B., Carberry, P., Hammer, G., Probert, M., Robertson, M., Holzworth, D., Huth, N., Hargreaves, J., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J., Silburn, M., Wang, E., Brown, S., Bristow, K., Asseng, S., Chapman, S., McCown, R., Freebairn, D., Smith, C., 2003. An overview of APSIM, a model designed for farming systems simulation. European Journal of Agronomy 18, 267–288.
Li, C., 2000. Modelling trace gas emissions from agricultural ecosystems. Nutrient Cycling in Agroecosystems 58, 259–276.
Li, C., Frolking, S., Frolking, T., 1992. A model of nitrous oxide evolution from soil driven by rainfall events: 1. Model structure and sensitivity. Journal of Geophysical Research: Atmospheres 97, 9759–9776.
Li, C., Frolking, S., Harriss, R., 1994. Modeling carbon biogeochemistry in agricultural soils. Global Biogeochemical Cycles 8, 237–254.
Low-Décarie, E., Chivers, C., Granados, M., 2014. Rising complexity and falling explanatory power in ecology. Frontiers in Ecology and the Environment 12, 412–418.
Myrgiotis, V., Rees, R., Topp, C., Williams, M., 2018. A systematic approach to identifying key parameters and processes in agroecosystem models. Ecological Modelling 368, 344–356.
Necpálová, M., Anex, R., Fienen, M., Del Grosso, S., Castellano, M., Sawyer, J., Iqbal, J., Pantoja, J., Barker, D., 2015. Understanding the DayCent model: calibration, sensitivity, and identifiability through inverse modeling. Environmental Modelling & Software 66, 110–130.
Oldfield, E., Bradford, M., Wood, S., 2019. Global meta-analysis of the relationship between soil organic matter and crop yields. SOIL 5, 15–32.
Parton, W., Holland, E., Del Grosso, S., Hartman, M., Martin, R., Mosier, A., Ojima, D., Schimel, D., 2001. Generalized model for NO_x and N_2O emissions from soils. Journal of Geophysical Research: Atmospheres 106, 17403–17419.
Pham, H., 2019. A new criterion for model selection. Mathematics 7, 1215. https://doi.org/10.3390/math7121215.
Robertson, G., Bruulsema, T., Gehl, R., Kanter, D., Mauzerall, D., Rotz, C., Williams, C., 2013. Nitrogen–climate interactions in US agriculture. Biogeochemistry 114, 41–70.
Sándor, R., Ehrhardt, F., Brilli, L., Carozzi, M., Recous, S., Smith, P., Snow, V., Soussana, J.-F., Dorich, C., Fuchs, K., Fitton, N., Gongadze, K., Klumpp, K., Liebig, M., Martin, R., Merbold, L., Newton, P., Rees, R., Rolinski, S., Bellocchi, G., 2018. The use of biogeochemical models to evaluate mitigation of greenhouse gas emissions from managed grasslands. Science of the Total Environment 642, 292–306.
Shahhosseini, M., Martinez-Feria, R., Hu, G., Archontoulis, S., 2019. Maize yield and nitrate loss prediction with machine learning algorithms. Environmental Research Letters 14, 124026. https://doi.org/10.1088/1748-9326/ab5268.
Thornton, A., Lee, P., 2000. Publication bias in meta-analysis: its causes and consequences. Journal of Clinical Epidemiology 53, 207–216.
Trenberth, K., 2011. Changes in precipitation with climate change. Climate Research 47, 123–138.
Vogeler, I., Giltrap, D., Cichota, R., 2013. Comparison of APSIM and DNDC simulations of nitrogen transformations and N_2O emissions. Science of the Total Environment 465, 147–155.
Weier, K., Doran, J., Walks, D., 1993. Denitrification and the dinitrogen/nitrous oxide ratio as affected by soil water, available carbon, and nitrate. Soil Science Society of America Journal 57, 66–72.
Weisburd, D., Piquero, A., 2008. How well do criminologists explain crime? Statistical modeling in published studies. Crime and Justice 37, 453–502.
Xu, Ri, Wang, Y., Wang, Y., Niu, H., Liu, Y., Zhuang, Q., 2019. Estimating N_2O emissions from soils under natural vegetation in China. Plant and Soil 434, 271–287.
Yeluripati, J., del Prado, A., Sanz-Cobeña, A., Rees, R., Li, C., Chadwick, D., Tilston, E., Topp, C., Cardenas, L., Ingraham, P., Gilhespy, S., Anthony, S., Vetter, S., Misselbrook, T., Salas, W., Smith, P., 2015. Global Research Alliance Modelling Platform (GRAMP): An open web platform for modelling greenhouse gas emissions from agro-ecosystems. Computers and Electronics in Agriculture 111, 112–120.
Yu, Y., Tao, H., Jia, H., Zhao, C., 2017. Impact of plastic mulching on nitrous oxide emissions in China's arid agricultural region under climate change conditions. Atmospheric Environment 158, 76–84.
Zhang, Y., Ma, M., Fang, H., Qin, D., Cheng, S., Yuan, W., 2020. Impacts of nitrogen addition on nitrous oxide emission: Comparison of five nitrous oxide modules or algorithms. Ecological Modelling 421, 108963. https://doi.org/10.1016/j.ecolmodel.2020.108963.
Zhang, Y., Niu, H., 2016. The development of the DNDC plant growth sub-model and the application of DNDC in agriculture: A review. Agriculture, Ecosystems and Environment 230, 271–282.
Zhang, Y., Yu, Q., 2020a. Identification of current research intensity and influence factors of agricultural nitrogen loss from cropping systems. Journal of Cleaner Production 276, 123308. https://doi.org/10.1016/j.jclepro.2020.123308.
Zhang, Y., Yu, Q., 2020b. What is the best article publishing strategy for early career scientists. Scientometrics 122, 397–408.
Zhao, J., Yang, Y., Zhang, K., Jeong, J., Zeng, Z., Zang, H., 2020. Does crop rotation yield more in China. A meta-analysis. Field Crops Research 245, 107659. https://doi.org/10.1016/j.fcr.2019.107659.
Zimmermann, J., Carolan, R., Forrestal, P., Harty, M., Lanigan, G., Richards, K., Roche, L., Whitfield, M., Jones, M., 2018. Assessing the performance of three frequently used biogeochemical models when simulating N_2O emissions from a range of soil types and fertiliser treatments. Geoderma 331, 53–69.