Evaluating the Potential of Legumes to Mitigate N_2O Emissions From Permanent Grassland Using Process-Based Models

Kathrin Fuchs[1,2], Lutz Merbold[1,3], Nina Buchmann[1], Gianni Bellocchi[4], Marco Bindi[5], Lorenzo Brilli[5,6], Richard T. Conant[7], Christopher D. Dorich[7], Fiona Ehrhardt[8], Nuala Fitton[9], Peter Grace[10], Katja Klumpp[4], Mark Liebig[11], Mark Lieffering[12], Raphaël Martin[4], Russell McAuliffe[12], Paul C. D. Newton[12], Robert M. Rees[13], Sylvie Recous[14], Pete Smith[9], Jean-François Soussana[8], Cairistiona F. E. Topp[13], and Val Snow[15]

[1]Department of Environmental Systems Science, Institute of Agricultural Sciences, ETH Zurich, Zurich, Switzerland, [2]Institute of Meteorology and Climate Research - Atmospheric Environmental Research, Karlsruhe Institute of Technology (KIT Campus Alpin), Garmisch-Partenkirchen, Germany, [3]Mazingira Centre, International Livestock Research Institute, Nairobi, Kenya, [4]UCA, INRAE, VetAgro Sup, Unité Mixte de Recherche sur l'Écosystème Prairial (UREP), Clermont-Ferrand, France, [5]DAGRI, University of Florence, Florence, Italy, [6]IBE-CNR, Florence, Italy, [7]Natural Resource Ecology Lab, Colorado State University, Fort Collins, CO, USA, [8]INRAE, Paris, France, [9]Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen, UK, [10]Institute for Future Environments, Queensland University of Technology, Brisbane, QLD, Australia, [11]NGPRL, USDA-ARS, Mandan, ND, USA, [12]AgResearch Grasslands Research Centre, Palmerston North, New Zealand, [13]Department of Agriculture, Horticulture and Engineering Sciences, Scotland's Rural College, Edinburgh, UK, [14]INRAE, Université de Reims Champagne-Ardenne, Unité Mixte de Recherche FARE, Reims, France, [15]AgResearch - Lincoln Research Centre, Christchurch, New Zealand

Abstract A potential strategy for mitigating nitrous oxide (N_2O) emissions from permanent grasslands is the partial substitution of fertilizer nitrogen (N_{fert}) with symbiotically fixed nitrogen (N_{symb}) from legumes. The input of N_{symb} reduces the energy costs of producing fertilizer and provides a supply of nitrogen (N) for plants that is more synchronous to plant demand than occasional fertilizer applications. Legumes have been promoted as a potential N_2O mitigation strategy for grasslands, but evidence to support their efficacy is limited, partly due to the difficulty in conducting experiments across the large range of potential combinations of legume proportions and fertilizer N inputs. These experimental constraints can be overcome by biogeochemical models that can vary legume-fertilizer combinations and subsequently aid the design of targeted experiments. Using two variants each of two biogeochemical models (APSIM and DayCent), we tested the N_2O mitigation potential and productivity of full factorial combinations of legume proportions and fertilizer rates for five temperate grassland sites across the globe. Both models showed that replacing fertilizer with legumes reduced N_2O emissions without reducing productivity across a broad range of legume-fertilizer combinations. Although the models were consistent with the relative changes of N_2O emissions compared to the baseline scenario (200 kg N ha^{-1} yr^{-1}; no legumes), they predicted different levels of absolute N_2O emissions and thus also of absolute N_2O emission reductions; both were greater in DayCent than in APSIM. We recommend confirming these results with experimental studies assessing the effect of clover proportions in the range 30–50% and $\leq$150 kg N ha^{-1} yr^{-1} input as these were identified as best-bet climate smart agricultural practices.

1. Introduction

Nitrous oxide (N_2O) is a powerful greenhouse gas (GHG) and plays a dominant role in stratospheric ozone depletion (Ravishankara et al., 2009; Smith et al., 2014). Agriculture contributes 58–84% to global anthropogenic N_2O emissions (Mosier et al., 2004; P. Smith et al., 2008) and is thus important for N_2O mitigation (Ciais et al., 2013; K. A. Smith, 2017). Most agricultural N_2O emissions arise from inputs of nitrogen (N) by being too high or out of synchrony with crop or pasture demand, resulting in the applied N being susceptible to losses through leaching or gaseous emissions including N_2O (Davidson & Kanter, 2014).

Grassland systems cover about 40% of the Earth's land surface and a quarter of these are located in temperate zones (White et al., 2000). Globally estimated N_2O emissions from managed grasslands contributed 54% to

the total agricultural emissions in 2006 (Dangal et al., 2019). These N_2O emissions significantly increased during recent decades due to increased mineral and organic fertilizer applications and excreta deposition from 1.74 Tg N_2O-N (1961) to 3.11 Tg N_2O-N (2014) (Dangal et al., 2019).

Besides fertilizer amendments (Flechard et al., 2007; K. A. Smith et al., 2012; Hörtnagl et al., 2018), recycling of N through the excreta of grazing animals (Saggar et al., 2004), aerial N deposition (Well & Butterbach-Bahl, 2010) and biological nitrogen fixation (BNF) via legume symbiosis with *Rhizobia* (Jensen et al., 2012) deliver important N inputs to grasslands while having the potential to increase N_2O emissions. Environmental variables, nitrogen rate, type of N input, timing, and placement of organic and mineral N fertilizers are important factors that influence the magnitude and duration of N_2O emission pulses (Baggs, 2011; Flessa, 2012; Shcherbak et al., 2014; K. A. Smith et al., 1997, 2012).

Nitrous oxide mitigation strategies for grasslands mainly focus on an increased N use efficiency (NUE) (K. A. Smith et al., 1997) by adjusting the amount, type, and timing of fertilizer to better match plant N demand (Bolan et al., 2004; Eckard et al., 2010; de Klein & Eckard, 2008; D. Li et al., 2013; Luo et al., 2010; Saggar et al., 2013). Other mitigation strategies aim to alter soil conditions affecting the microbial production of N_2O. Approaches here include liming of acid soils and the application of nitrification inhibitors (Cahalan et al., 2015; Galbally et al., 2010; Lam et al., 2017). The use of legumes has also been proposed as a mitigation strategy (Jensen et al., 2012; Rochette & Janzen, 2005). Legumes can (a) use symbiotically fixed N (N_{symb}) from the atmosphere as substitute for mineral fertilizer, which has high GHG costs associated with its manufacture (1.6–6.4 kg CO_2-eq per kg fertilizer N; Brentrup & Pallière, 2008) and (b) provide a N source for plants that is better synchronized to plant demand than the alternative, which is episodic applications of N fertilizer. In addition, there is evidence that suggest mixtures of legumes and grasses can produce higher yields than monocultures of grasses or legumes ("overyielding") (Finn et al., 2013; Kirwan et al., 2007; Lüscher et al., 2014). The potential to sustain or even increase yields would be a highly desirable characteristic of any mitigation option both in terms of adoption of the technology by farmers (Vellinga et al., 2011) and its benefit in reducing GHG emissions per area of land and per unit of product, referred to as "N_2O emission intensity" (N_2O-Int) (van Groenigen et al., 2010).

Despite its theoretical advantages, the potential to mitigate N_2O emissions with legumes remains largely untested (Flessa et al., 2014; Osterburg et al., 2013). Only few field experiments have measured N_2O emissions for different legume proportions in permanent grassland systems, and these have shown mixed results (Ammann et al., 2009; Fuchs et al., 2018; Hörtnagl et al., 2018; Ingram et al., 2015; Jensen et al., 2012; Klumpp et al., 2011; D. Li et al., 2011; Schmeer et al., 2014; Simek, 2004; Virkajärvi et al., 2010) (Figure 1; see Supporting Information Text S1). Although N_2O emissions across these studies were highly variable, a smoothed fit suggests constant emissions across the full range of clover proportions used in these studies (Figure 1b). In contrast, N_2O emissions increase with increased fertilizer rates (Figure 1c). In order to test the robustness of legumes as a N_2O mitigation option, it would be valuable to extend the experimental evidence to better cover a wider range of legume proportions, for example, by examining the potentially interesting combination of a low N fertilizer level and relatively high (30–80%) proportion of clover (Figure 1).

Currently, the potential of grass-legume mixtures is poorly exploited (Phelan et al., 2015). While New Zealand and Switzerland have promoted grass-legume mixtures and have grasslands with up to 30% legume content as the norm, the use of legumes is not the standard in many other countries. The area of grasslands containing legumes is furthermore not routinely captured in country statistics nor are the legume proportions. Similarly, there are large spatial and temporal variations in legume proportions caused by varying seed mixes and difficulties in legume establishment, management, and persistence. Due to this lack of data, a quantitative estimate of the *status quo* is difficult. This motivated the assessment of several simplified scenarios to approximate the magnitude of the global mitigation potential.

Costs and workload constrain experimental measurements to a few locations and seasons, with restrictions on the number of treatments and replicates. Overall, experiments investigating N_2O emissions in swards of different clover proportions are limited by the temporal coverage and/or by the number of treatments (clover-fertilizer combinations). Studies applying the eddy covariance (EC) technique to acquire data of high temporal coverage and resolution can clearly resolve N_2O emission peaks at the field-scale (Fuchs et al., 2018; Hörtnagl et al., 2018), while the number of simultaneously observed treatments in EC studies is very limited.
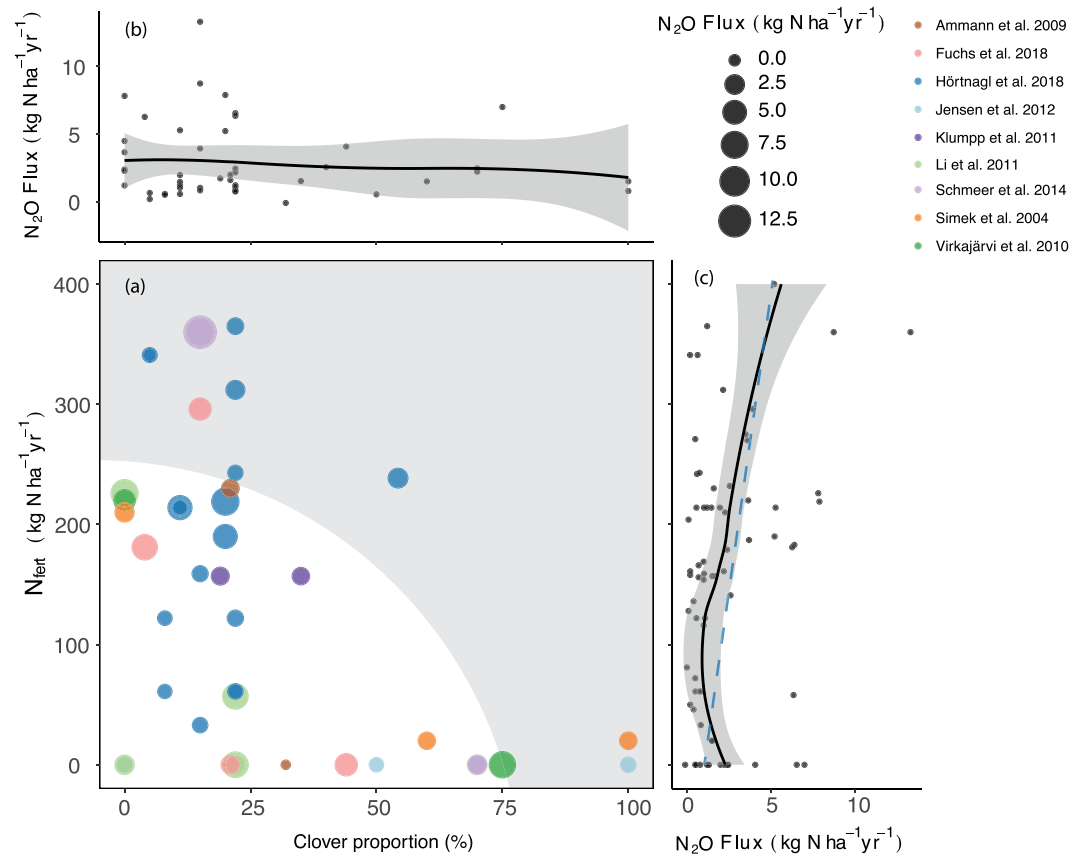
Figure 1. Observed data from literature (source = color) with documented N_2O emissions, clover proportions and fertilizer amounts. Panel (a) relates observed clover proportions (horizontal axis) and fertilizer levels (vertical axis) at different sites to N_2O emissions (bubble area). The gray-shaded area in (a) indicates that these combinations are unrealistic or of little or no practical interest, but some are important for testing model performance. The plots at the margins relate (b) clover proportion to N_2O emissions across all fertilizer levels and (c) fertilizer amendments to N_2O emissions across all clover proportions. The dashed blue line indicates N_2O emissions according to the IPCC emission factor (1% of N_{fert}), black lines indicate locally weighted regression (loess) fits of the displayed data and the shaded area their 95% confidence intervals (Cleveland & Devlin, 1988). All studies are from the temperate zone except Virkajärvi et al. (2010), which was in boreal grasslands characterized by freeze-thaw cycles.

Studies with more treatments tend to use manual GHG chambers to investigate N_2O emissions, but these are generally limited by their coarse temporal resolution (e.g., Niklaus et al., 2016) and are therefore less capable of resolving temporal dynamics. However, if sampling frequency is too low, data are less suitable for annual N_2O flux budgets (Barton et al., 2015). On the other hand, manual and automatic GHG chamber studies with higher sampling frequency (Klumpp et al., 2011; D. Li et al., 2011) covering a full growing season have often been constrained to two to five clover-fertilizer combinations only. In addition, chambers cover relatively small sampling areas (usually less than a few m^2). While any in situ measurement delivers valuable data, a systematic assessment of the effect of different clover proportions on N_2O emissions at different sites and scales (plot to ecosystem to catchment) is clearly lacking and unlikely to be resolved experimentally.

Biogeochemical process models allow for the ex-ante and ex-post estimation of GHG emissions (including N_2O emissions) under many different management regimes. In particular, they can be used for a systematic assessment of mitigation strategies across several sites, investigating different rates of N fertilizer and levels of legume proportion simultaneously, which has so far not been achieved in field N_2O measurements. Models can predict these different scenarios over a long time-span and reveal important long-term effects (e.g., Lugato et al., 2018).

In this study, we use the biogeochemical models APSIM (Holzworth et al., 2014) and DayCent (Del Grosso et al., 2001; Parton et al., 1998, 2001). These models have the ability to simulate key processes of C and N
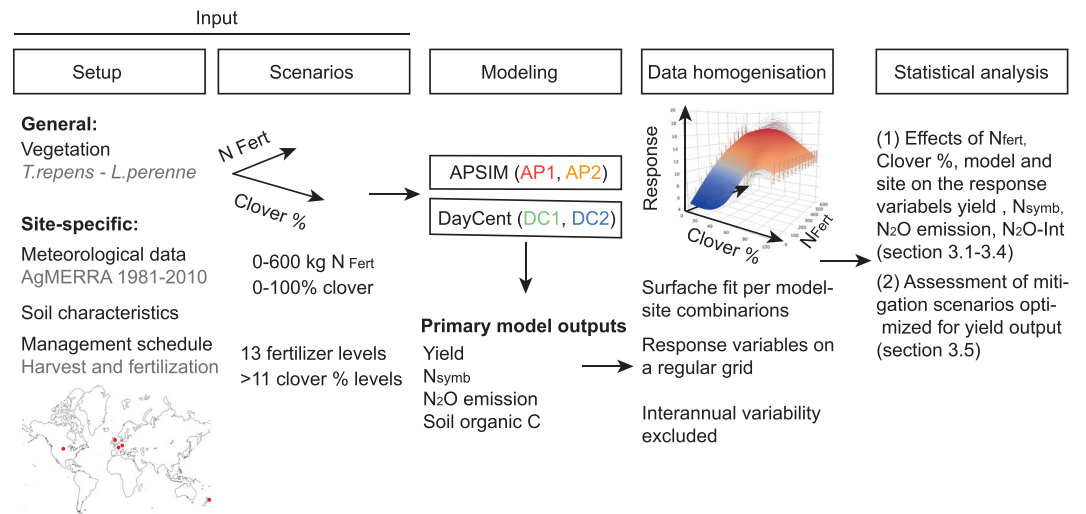
Figure 2. Overview of the modeled scenarios: 13 levels of varying fertilizer amount (0–600 kg N ha^{-1} yr^{-1}) and >11 levels of clover proportion (0–100%) were simulated. Fertilizer types included both organic fertilizer (OF) and mineral fertilizer (MF). All simulations were run at five sites G1 (USA), G2 (New Zealand), G3 (France), G4 (United Kingdom), and G5 (Switzerland) by the biogeochemical process models APSIM and DayCent, both in two variants (see text for details). In order to obtain the response variables on a homogeneous grid and exclude inter-annual variability, surfaces were fitted to model outputs (yield, N$_2$O emissions) per model-site combination on the surface expanded by actual clover percentage and annual fertilizer amount.

cycling, including key processes of N$_2$O production, nitrification, and denitrification and, importantly, have been validated with field data at the study sites used with respect to N$_{symb}$ (Fitton et al., 2019) and combinations of biomass yields and N$_2$O emissions (Ehrhardt et al., 2018; Fuchs et al., 2020). The number of deployed models in this current study was lower than those in the publication cited above because most of the biogeochemical models were not able to simulate grass-legume interaction explicitly, as required for this study.

We conducted a simulation study at five temperate grassland sites using the two biogeochemical models, each in two variants. Our specific objectives were (1) to systematically assess the effects of a wide range of legume proportions and N fertilizer amounts on yields, N$_2$O emissions, and N$_2$O emission intensities (N$_2$O-Int) and (2) to identify the optimal N$_2$O mitigation strategy.

2. Materials and Methods

A set of theoretical scenarios, that is, combinations of legume proportions (0–100%) and N fertilizer amounts (N$_{fert}$; 13 levels 0–600 kg N ha^{-1} yr^{-1}) in a full factorial design, was run at five temperate grassland sites across the globe for a simulation period of 30 years (1981–2010) (Figure 2). Although the higher fertilizer rates are of little practical relevance, they are important for testing model responses and constraining statistically derived surfaces. Grass-clover mixtures of two species *Lolium perenne* and *Trifolium repens* were modeled as examples of species that are widespread across the temperate climatic zone. Mineral fertilizer in form of ammonium nitrate was simulated. The starting point for this study was the final stage of an international modeling intercomparison exercise initiated by the Global Research Alliance (Integrative Research Group) as described by Ehrhardt et al. (2018). For simplicity, and in order to exclude the confounding effects of grazing by further enhancing variability, we performed the study on non-grazed grassland systems. Mitigation of N$_2$O emissions via changes in grazing practices (i.e., livestock density and N fertilizer inputs) have previously been investigated by Sándor et al. (2018).

2.1. Models Used in This Study

This exercise was performed by a consortium of modelers with the biogeochemical models APSIM (Holzworth et al., 2014; in two variants) and DayCent (Del Grosso et al., 2001; Parton et al., 1998; in two variants) according to an agreed protocol. Models were previously calibrated (corresponding to Stage 5 in

Ehrhardt et al., 2018) with respect to plant productivity, soil moisture and temperature, soil mineral N content, and model uncertainties of the respective models were evaluated in an international model intercomparison (Ehrhardt et al., 2018; Table S1) with further testing of yield and N_2O emissions at two levels of pasture clover content in Fuchs et al. (2020). Biological nitrogen fixation and yields in grass-clover mixtures were validated at one site described by Fitton et al. (2019). The scenarios required modelers to manipulate the legume proportion of the grasslands. In a physical experiment, this would be achieved by sowing or over-sowing the desired species, with the ultimate legume proportion resulting from interactions between management and environmental conditions (see Fuchs et al., 2018). APSIM simulates grasslands in a way that is comparable with a physical experiment where the modeler can influence, but not directly control, the proportion of legumes (dynamic legume proportion), while in DayCent the legume proportion is explicitly specified by the modeler (Table S2; Fitton et al., 2019). Biological nitrogen fixation from grasslands with varying legume proportion can be simulated by APSIM directly as the dependency of BNF on species composition is represented in the model (Text S2). In contrast, DayCent requires the user to specify a potential maximum N_{symb} that should, sensibly, vary with legume proportion. In Fitton et al. (2019), we developed a method to derive this DayCent parameter (potential maximum N_{symb}) based on the legume percentage of the pasture. Note that DayCent internally reduces the actual N_{symb} as fertilization increases. This method is described in Text S3 and was validated in Fitton et al. (2019).

2.1.1. APSIM

APSIM (v7.10 r4162), short for Agricultural Production Systems sIMulator (Holzworth et al., 2014), is a process-oriented simulation framework comprising several modules. The model primarily operates on a daily time-step, but several individual modules work with shorter internal time discretization. APSIM was used for this study in two variants differing in their soil water modules, the SWIM module based on Richards' equation (here named AP1; Huth et al., 2012; SWIM for Soil and Water integrated Model), and the capacitance-based Soil Water module (here named AP2; Probert et al., 1998; further details can be seen in Text S2). Sward growth was simulated with the AgPasture module (F. Li et al., 2011) with recent additions to model reallocation of N reserves (Vogeler & Cichota, 2016) and an implementation of the Penman-Monteith equation designed for intermingled canopies (Snow & Huth, 2004). Soil organic matter and nitrogen transformations were simulated with the SoilN module (Probert et al., 1998) with a derivation of the $N_2:N_2O$ ratio adapted from DayCent (Thorburn et al., 2010) and a simple routine to calculate volatilization losses (Vogeler et al., 2010). APSIM simulates N_2O production from nitrification and denitrification. Nitrification is implemented with a Michaelis-Menten approach, with a fixed fraction of NH_4^+ for nitrification. N_2O production via nitrification is down-regulated at suboptimal temperature, moisture, and pH in the respective layer. Denitrification depends on NO_3^- and active C concentrations in the respective layer. The maximum denitrification rate is regulated downwards and depends on temperature and soil moisture in each layer (for further details on the simulation and parameterization of BNF and N_2O emissions in APSIM, see Text S2). Simulated pasture composition in APSIM is dynamic and adapts to prevailing conditions, such as fertilizer inputs, much as one might expect a physical pasture to behave. To create variation in legume content for any given fertilizer regime, the simulated pastures were resown every third year, much as might be done in a physical experiment.

2.1.2. DayCent

DayCent, the daily time-step model that has evolved from the Century model (Del Grosso et al., 2001; Parton et al., 1998), was used in this study in the variants DayCent v4.5 2010 (here named DC1) and DayCent v4.5 2013 (here named DC2). These variants differ in their calculations of solar radiation and in their calculations of maintenance and growth respiration, as well as the implementation of N_2O emissions due to freeze-thaw cycles in the latter (please find further details in Text S3). DayCent includes four major sub-models: (1) the plant growth sub-model calculates biomass production and allocates net primary production (NPP) to various plant pools; (2) the soil organic matter (SOM) sub-model simulates decomposition of dead plant material (litter) and SOM and allocates soil carbon to three SOC pools and the litter pool; (3) the soil water sub-model simulates water flow between different layers; and (4) the trace gas flux sub-model simulates gaseous emissions. DayCent simulates N_2O losses based on the leaky pipe metaphor where nitrification and denitrification are conceptualized as leaky pipes through which N is cycling but losing some gaseous N through a hole (Parton et al., 2001). Nitrification and denitrification rates depend on NH_4^+ (for nitrification) and NO_3^- concentrations and labile carbon availability (for denitrification), soil water content and

temperature. While the proportion of the nitrification and denitrification escaping as gaseous N as either (NO_x, N_2O, or N_2) is fixed, the ratio of N_2O with respect to NO_x and N_2 depends on soil water content and soil properties regulating soil gas diffusivity (Parton et al., 2001; see Text S3).

2.1.3. Empirical Findings for the Effect of N_{fert} and Clover on Yields

In order to present model results in the context of empirical findings, we displayed example data from two studies for the dependency of yields on fertilizer N (Reid, 1983; here E1) and clover proportion (Nyfeler et al., 2009; here E2) in Figure 4. These can be regarded as case studies with the purpose of showing the reader what patterns we would expect. We used the regression equation from Reid (1983) as displayed in their Table 2. In contrast to our model outputs, grass-clover mixtures in Reid (1983) did not contain exactly 50% clover in dry biomass. Besides Reid (1983), see also further studies of Sparrow (1979), Prins (1983), and Whitehead (1995) for comparable data. To provide empirical evidence for the effect of clover percentage on yields, we used the equation from Nyfeler et al. (2009), equation 1 with the coefficients according to their Table S3 for mixtures of *T. repens* and *L. perenne*, excluding the other species. The Nyfeler study had 50 kg N ha^{-1} yr^{-1} as the minimum fertilizer rate instead of 0 kg N ha^{-1} yr^{-1} as in our study.

2.2. Site Information

The sites used for the simulations were five grassland sites in the temperate zone (Table S3). Daily weather data for 1980–2010 was supplied by A. Ruane from the AgMERRA data set (Ruane et al., 2015). Each modeling group used their plant growth model parameterizations, with the exceptions noted below, as developed in the fully informed or calibrated (Stage 5) modeling reported by Ehrhardt et al. (2018). The exception to the parameterization was in relation to the grass-legume ratio and symbiotic N fixation as described in the specific model sections in the supporting information.

2.3. Validation Results

For net primary production, the applied models in the validation exercise showed relative biases of −0.04– 0.74 and RRMSE of 0.70–0.89 (Ehrhardt et al., 2018). For N_2O emissions, relative biases were −0.80–0.03 and RRMSE ranged between 0.13 and 0.81 (Ehrhardt et al., 2018; see Table S1a). A detailed validation of N_2O emissions and its drivers at the Swiss site revealed that both models used here generally performed better than the Swiss IPCC calculations for that site (Fuchs et al., 2020).

The relationship between yields and legume percentage was validated in Fitton et al. (2019), showing that APSIM most accurately predicted yields at 0% and 100% clover (bias below 2%), but underestimated yields at 25–50% clover by 23–27% and overestimated yields at 75% clover by 10%. DayCent estimated yields with a bias below 10% across all clover percentages, except at 75% clover percentage where it overestimated yields by 33% (Fitton et al., 2019; see Table S1b).

2.4. Modeling Protocol

Management events included cutting and fertilization. Cutting events were simulated monthly during the growing season (May-October for the European sites G3–5, May-September for G1 in North Dakota, and September-May for G2 in New Zealand; each on the first day of the month) and followed by fertilization on the fifth day of the month (Table S4). In line with agricultural practices, the exception deviating from this scheme was that there was no fertilizer applied after the last harvest of the season and fertilization in spring started in the month before the first cut in each year (Table S4). The residual biomass after all cuts was 1,200 kg DM ha^{-1}, and all cut biomass was removed from the simulation. The total amount of N to be applied annually was split equally between the individual applications, which were surface applied. The mineral N was applied as ammonium-nitrate (NH_4-NO_3). Aerial deposition of N was set to zero. The sites were presumed to be flat; hence, there were no N losses via run-off. Model outputs included yields, harvested C and N, plant N uptake, N_2O emissions, soil organic carbon (SOC), soil organic nitrogen (SON), nitrate and ammonium concentrations, and nitrate leaching (Table S5).

2.5. Statistical Analysis

The dynamic simulation of clover proportions in APSIM resulted in clover proportions that were irregularly distributed between 0% and 100% and therefore not suitable for direct analysis. From visual inspection, we observed no trends in the output data. In order to (1) obtain a response value on a regular grid for any clover-fertilizer combination per site, per model and per fertilizer type and to (2) exclude the inter-annual

variability and analyze an aggregated outcome across years, we fitted response surfaces on the area spanned by clover proportion and fertilizer amount (Figure 2; see also Figure S1 for an example surface fit). These smooth surfaces can be considered as an interpolation, aggregating the 30 years of model outputs. We used the tensor product smoother to include two dimensions (clover proportion and fertilizer amount), suitable due to its scale-invariant properties (Wood, 2006). We prescribed a fixed number of knots for both dimensions (k = 4, resulting in $4^2 = 16$ knots) in order to hold the number of parameters consistent across all surface fits. We had also investigated surface responses using a linear model with second order polynomials for both predictors and an interaction. However, residual analysis of the results from that approach showed that the polynomial fits violated the assumptions of the linear model (i.e., residual distribution showed no normality and homogeneous variance) in most site-model combinations. Thus, from visual inspection and residual analysis, the chosen method using the tensor product smoother better represented the characteristics of the data and was therefore chosen for this study. Daily model outputs were aggregated to annual values before the fitting. The surface fitting and statistical analyses were carried out using the open source software R (R Core Team, 2016) and specifically the "mgcv" package (Wood, 2011) for fitting smoothing splines, building the tensor product of the predictors (N_{fert} and clover %) with the function te(), and fitting the smoothing splines on this tensor product with gam() (Wood, 2011). Further, the "data.table" package (Dowle & Srinivasan, 2017) was used for efficiently handling the large data sets. ANOVA was used in order to systematically asses the effects of clover proportion, fertilizer amount, model and site on yields, N_{symb}, N_2O emissions, and N_2O emission intensities (N_2O-Int calculated as mass unit of N_2O emissions per mass unit biomass yield). For the statistical analysis, the post-processed data set (after surface fitting) was used as this had the advantage of a balanced design. The ANOVA was performed separately for APSIM and DayCent, since the patterns varied between models such that averaging across model types would make it impossible to address model-specific behaviors and their implications on the results. To compare the effect of clover proportion and fertilizer amount on a response variable (e.g., yield, N_2O emissions, and N_2O-Int) across several sites, the response variable was presented as standardized response. Here, the standardization was consistently defined as the response variable divided by the maximum of the respective response variable across all scenarios (combinations of clover proportion and fertilizer amount) per site and model.

$$Y'_{i,m,s} = \frac{Y_{i,m,s}}{\max \{Y_{i,m,s}\}}$$

$Y'_{i,m,s}$ is the standardized response variable for scenario i, fixed model m, and fixed site s, and $Y_{i,m,s}$ is the original response variable. The results are presented per model (DayCent, APSIM) across model variants in order to assess the general pattern rather than the individual model variants' differences. Mean values and ranges given in the text and figures refer to the mean and range across variants and not the inter-annual which was excluded by the surface fitting procedure.

A baseline scenario of no clover with 200 kg N ha^{-1} yr^{-1} fertilizer applied was used for comparing mitigation scenarios. This baseline reasonably represents intensively used cut-and-carry systems, corresponding to N recommendations for fertilized grassland present in several countries. To test if a small clover percentage would significantly change the results under the baseline, we regarded clover percentages of 5% as a second baseline scenario. Here, we were not trying to mimic the actual management at the sites nor intending to address site-specific optima (species, N rates, grazing). Instead, we conceptually prescribed fixed management schedules in order to look at the response surfaces of fixed comparable setups across sites. While corresponding approximately to actual fertilizer amounts at the G5 site, which was the highest input system of any of the sites in this study, this baseline deviated from the actual fertilizer amendments on the other sites (Table S3). Nevertheless, we used this as baseline because we were interested in the general patterns using the sites as "generalizable" examples.

2.6. Estimating the Global Technical Mitigation Potential of Replacing N_{fert} With BNF

Our work focuses on managed pasturelands with C3 grasses and is not applicable to extensive rangeland and C4 pastureland. Dangal et al. (2019) estimated that pastures based on C3 grasses cover 570 million ha of the terrestrial surface, based on the HYDE3.2 data set by Klein Goldewijk et al. (2017). We used the estimated global amount of 2.9 Tg N_2O-N yr^{-1} for the recent period 2001–2014 together with their estimate of 11% of these emissions being attributed to synthetic fertilizer, resulting in emissions of 319 Gg Tg N_2O-N yr^{-1}

Table 1
Effects of Site, Model, Fertilizer Amount, and Clover Percentage on Yield, N_2O Emission, and N_2O-Int

		Yield t DM ha^{-1} yr^{-1}		N_2O emission kg N_2O-N ha^{-1} yr^{-1}		N_2O-Int g N_2O-N kg^{-1} DM^{-1}	
		APSIM	DayCent	APSIM	DayCent	APSIM	DayCent
	Grand mean	10.3	6.8	0.52	5.31	0.04	0.83
Site		***	***	***	***	***	***
	G1	8.3[a]	5.0[a]	0.38[b]	2.53[a]	0.034[a]	0.595[a]
	G2	11.9[b]	7.3[b]	0.53[cd]	3.57[b]	0.034[a]	0.494[a]
	G3	10.7[c]	6.7[bc]	0.43[a]	7.11[c]	0.028[a]	1.071[b]
	G4	9.2[b]	6.3[c]	0.61[de]	6.32[c]	0.054[b]	1.015[b]
	G5	11.3[bc]	8.8[d]	0.78[e]	10.4[d]	0.059[b]	1.231[b]
Model		***	***	ns	***	***	***
	AP1	10.8[a]	-	0.55[a]	-	0.037[a]	-
	AP2	9.8[b]	-	0.51[a]	-	0.043[b]	-
	DC1	-	5.8[a]	-	4.91[a]	-	0.917[a]
	DC2	-	7.8[b]	-	5.75[b]	-	0.751[b]
N_{fert}		***	***	***	***	***	***
	0	5.3[a]	5.8[a]	0.01[a]	1.83[a]	0.006[a]	0.341[a]
	150	9.1[b]	6.7[b]	0.31[b]	4.35[b]	0.025[b]	0.681[a]
	300	11.3[c]	7.0[b]	0.66[c]	6.59[c]	0.054[c]	0.989[b]
	450	12.4[cd]	7.1[b]	1.21[d]	8.37[cd]	0.096[d]	1.232[bc]
	600	13.3[d]	7.3[b]	1.90[e]	9.62[d]	0.145[e]	1.392[c]
Clover %		***	***	***	ns	***	ns
	0	8.7[a]	6.1[a]	0.50[a]	4.80[a]	0.044[a]	0.886[a]
	25	9.5[a]	7.1[bc]	0.48[a]	5.31[a]	0.040[ab]	0.785[a]
	50	10.7[b]	7.3[c]	0.46[a]	5.51[a]	0.034[b]	0.769[a]
	75	11.8[b]	6.9[bc]	0.54[ab]	5.51[a]	0.034[b]	0.823[a]
	100	10.8[b]	6.5[ab]	0.68[b]	5.46[a]	0.050[ab]	0.884[a]

Note. Predicted values per model (APSIM, DayCent) are shown for mineral fertilizer scenarios (see Table S7 including organic fertilization). Different small letters indicate significant differences among levels (Tukey's HSD) per factor (Site, Model, N_{fert}, Clover %). ns indicates no significant differences for the respective factor.
***Significant effect at a significance level of $p < 0.01$.

attributed to fertilizer application during this period. Based on this estimate of N_2O emissions from global fertilizer application on grasslands, we used our estimated relative values for N_2O mitigation averaged across sites to calculate the global technical mitigation potential of replacing N_{fert} with BNF under the baseline and different mitigation scenarios. For the global estimate, the mitigation effect was considered relative to the N_2O emissions attributed to fertilizer use, which are the site emissions at the fertilizer scenario minus the background N_2O emissions, that is, N_2O emissions under the scenario of no fertilizer nor symbiotically fixed N input.

3. Results

3.1. Nitrous Oxide Emissions

Nitrous oxide emissions ranged between 0.01 and 1.9 kg N_2O-N ha^{-1} yr^{-1} (0.1–0.4% of N_{fert}) across scenarios for APSIM variants and between 1.8 and 9.6 kg N_2O-N ha^{-1} yr^{-1} (1.4–5.5% of N_{fert}) for DayCent model variants averaged across sites (Table 1, Figure 3). The emissions varied significantly among models, sites, and fertilizer levels (Tables 1 and S4). Clover proportion did not affect N_2O emissions in DayCent (Table 1) and did not affect N_2O emissions in APSIM at 0–75% clover but significantly increased N_2O emissions at the 75–100% clover level (Table 1) due to increased N_2O emissions at high fertilizer levels of 450–600 kg N ha^{-1} yr^{-1} in clover-rich grasslands (Table 2, Figure 3). The N_2O emissions depended not only on the amount of N input but also on the N source. For the same total N input (here defined as $N_{fert} + N_{symb}$), lower N_2O emissions were found when a larger fraction of the N input came from symbiotic N fixation (as N_{symb}) rather than from fertilization (N_{fert}). The mitigation option would be ineffective if N_2O emissions were independent of the source of N; thus, this finding shows that an important precondition for effectiveness of the mitigation strategy was met. The N_2O emissions were highest at maximal fertilizer levels and
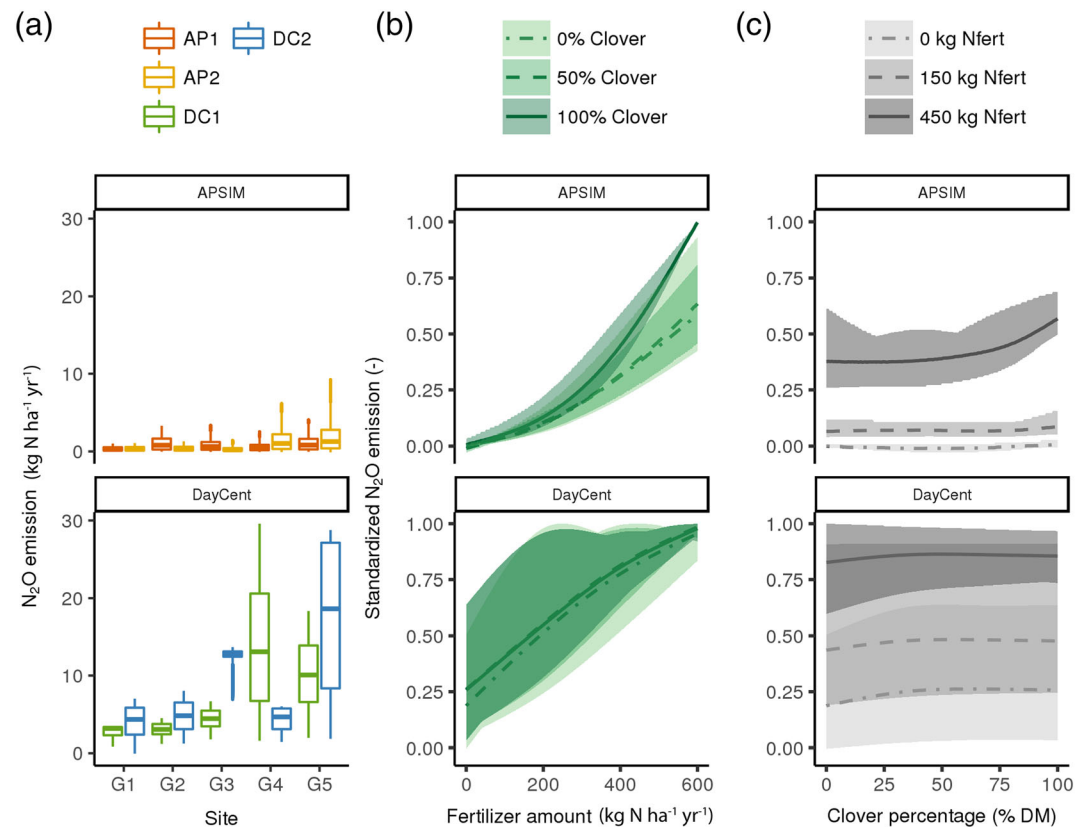
Figure 3. (a) Boxplots of N_2O emissions across all mineral N scenarios (clover-fertilization combinations) per site (G1–G5) for APSIM (top) and DayCent variants (bottom); (b) dependency of standardized N_2O emissions (y-axis) on fertilizer amount (x-axis) at 0% clover percentage (dot-dashed) 50% clover percentage (dashed) and 100% clover percentage (solid line); and (c) dependency of standardized N_2O emissions (y-axis) on clover percentage (x-axis) for the fertilizer levels 0 kg N (dot-dashed), 150 kg N (dashed), and 450 kg N (solid line) (APSIM: top, DayCent: bottom). N_2O emissions in b and c were standardized with the maximum yields per model-site combination. Lines display the mean across sites, shaded areas indicate the range across sites per model family.

were generally smallest in the absence of fertilizer. While N_2O emissions were higher for the DayCent variants, APSIM variants predicted higher losses via nitrate leaching (Figure S4).

3.1.1. Effect of Fertilization on Nitrous Oxide Emissions

In the absence of fertilization, models simulated average N_2O emissions of 0.01 kg N_2O-N ha^{-1} yr^{-1} (APSIM) and 1.83 kg N_2O-N ha^{-1} yr^{-1} (DayCent) across clover proportions (Table 1). At 150 kg N fertilization levels, N_2O emissions increased to 0.31 kg N_2O-N ha^{-1} yr^{-1} (APSIM) and 4.35 kg N_2O-N ha^{-1} yr^{-1} (DayCent) (Table 2). At maximal fertilization levels (600 kg N), N_2O emissions reached 1.9 kg N_2O-N N ha^{-1} yr^{-1} (APSIM) and 9.6 kg N_2O-N ha^{-1} yr^{-1} (DayCent), averaged across clover scenarios, model variants and sites (Table 1). Fertilization increased N_2O emissions for all model-site-combinations. However, different types of relationships were observed. The APSIM models showed an almost exponential increase of N_2O with fertilization, (Figure 3), while the DayCent models simulated approximately linear relationships, and in some cases (e.g., G3) lower N_2O emission increments at higher fertilizer rates (leveling-off) (Figure 3).

3.1.2. Effect of Clover Proportion on Nitrous Oxide Emissions

In general, clover proportion did not affect N_2O emissions across models and sites (Figure 3, Tables 1 and 2) between 0% and 75% clover (Table 1). As an exception, APSIM simulated significantly increased N_2O emissions at a clover proportion of 100% in the higher fertilized scenarios ($\geq$450 kg $N_{fert.}$), depicting the effects of high N inputs in these scenarios. However, one has to keep in mind that those clover-fertilizer combinations are rather artificial and of no practical importance. A notable characteristic was that DayCent simulated background N_2O emissions of 1.8 (0.4–4.3) kg N_2O-N ha^{-1} yr^{-1}, without N fertilization or clover in the field. APSIM showed no N_2O fluxes without N fertilization independent of clover level.

Table 2

Average and Range of Yields, N$_2$O Emission, and N$_2$O-Int for Different Combinations of N$_{fert}$ and Clover Proportion Across Sites and Model Variants for APSIM and DayCent for Mineral Fertilizer N Scenarios

	N 0	N 150	N 300	N 450	N 600
APSIM Yield (t DM ha^{-1} yr^{-1})					
0% clover	2.0 (0.9–3.6)	6.5 (5.4–7.6)	9.7 (8.7–11.2)	11.7 (9.7–13.8)	13.4 (9.7–16.0)
25% clover	2.2 (1.2–3.4)	7.2 (6.1–8.7)	10.8 (9.4–12.1)	12.9 (10.9–14.1)	14.2 (11.2–15.8)
50% clover	3.9 (2.6–5.3)	9.1 (7.5–10.9)	12.2 (10.6–13.7)	13.6 (11.6–15.0)	14.5 (16.4–15.7)
75% clover	7.5 (4.9–9.9)	11.8 (10.1–13.4)	13.1 (10.9–14.8)	13.1 (10.5–14.6)	13.6 (10.5–15.7)
100% clover	11.1 (8.5–13.3)	10.9 (7.3–13)	10.7 (7.0–12.9)	10.7 (7.0–12.9)	10.7 (7.0–12.8)
DayCent Yield (t DM ha^{-1} yr^{-1})					
0% clover	3.2 (1.6–4.2)	5.8 (4.9–7.3)	6.8 (5.8–8.6)	7.0 (5.5–8.9)	7.6 (5.8–9.7)
25% clover	5.9 (3.7–7.6)	7.0 (5.2–8.9)	7.4 (5.6–9.4)	7.5 (5.4–9.5)	7.7 (5.5–9.9)
50% clover	7.0 (4.8–9.1)	7.3 (5.2–4.9)	7.4 (5.3–9.5)	7.5 (5.3–9.6)	7.5 (5.3–9.7)
75% clover	6.9 (5.0–9.0)	6.9 (5.1–9.0)	7.0 (5.1–0.1)	7.0 (5.1–9.1)	7.1 (5.1–9.2)
100% clover	6.5 (4.6–8.6)	6.5 (4.9–8.6)	6.5 (4.9–8.6)	6.5 (4.9–8.5)	6.5 (4.9–8.5)
APSIM N$_2$O emission (kg N$_2$O-N ha^{-1} yr^{-1})					
0% clover	0 (0–0)	0.2 (0.1–0.3)	0.6 (0.3–1.0)	1.1 (0.6–1.9)	1.7 (0.9–3.1)
25% clover	0 (0–0)	0.2 (0.10.3)	0.6 (0.3–1.0)	1.1 (0.5–2.0)	1.8 (0.8–3.2)
50% clover	0 (0–0)	0.2 (0.1–0.3)	0.6 (0.2–1.0)	1.2 (0.5–2.1)	2.0 (0.8–3.5)
75% clover	0 (0–0)	0.2 (0.1–0.3)	0.6 (0.3–1.1)	1.3 (0.6–2.4)	2.3 (1.0–4.3)
100% clover	0 (0–0)	0.2 (0.1–0.4)	0.8 (0.4–1.5)	1.8 (0.7–3.6)	3.4 (1.0–6.7)
DayCent N$_2$O emission (kg N$_2$O-N ha^{-1} yr^{-1})					
0% clover	1.9 (0.4–4.3)	4.6 (2.3–7.6)	7.5 (3.7–13.5)	10.0 (4.7–19.8)	11.9 (5.1–22.9)
25% clover	2.4 (0.5–5.3)	5.0 (2.3–7.9)	7.9 (3.8–14.4)	10.4 (4.6–20.6)	12.2 (5.1–23.4)
50% clover	2.6 (0.6–5.6)	5.2 (2.4–8.0)	8.1 (3.8–14.6)	10.6 (4.6–20.7)	12.4 (5.1–23.5)
75% clover	2.6 (0.6–5.6)	5.2 (2.4–8.0)	8.0 (3.7–14.4)	10.6 (4.6–20.5)	12.4 (5.1–23.3)
100% clover	2.6 (0.6–5.6)	5.1 (2.3–7.9)	8.0 (3.7–14.3)	10.6 (4.6–20.3)	12.4 (5.1–23.0)
APSIM N$_2$O-Int (g N$_2$O-N kg^{-1} DM^{-1})					
0% clover	0.010 (0.003–0.034)	0.028 (0.019–0.043)	0.059 (0.040–0.096)	0.097 (0.061–0.159)	0.134 (0.079–0.220)
25% clover	0.006 (0.001–0.018)	0.027 (0.018–0.04)	0.057 (0.033–0.090)	0.093 (0.051–0.149)	0.131 (0.069–0.210)
50% clover	0.002 (0.000–0.006)	0.023 (0.013–0.034)	0.052 (0.025–0.082)	0.092 (0.045–0.149)	0.140 (0.073–0.228)
75% clover	0.001 (0.001–0.002)	0.017 (0.010–0.026)	0.049 (0.03–0.079)	0.105 (0.063–0.177)	0.179 (0.108–0.313)
100% clover	0.001 (0.001–0.006)	0.025 (0.016–0.037)	0.075 (0.048–0.128)	0.172 (0.108–0.315)	0.311 (0.166–0.585)
DayCent N$_2$O-Int (g N$_2$O-N kg^{-1} DM^{-1})					
0% clover	0.63 (0.41–1.05)	0.82 (0.45–1.18)	1.09 (0.55–1.57)	1.34 (0.70–2.12)	1.49 (0.80–2.35)
25% clover	0.42 (0.26–0.84)	0.73 (0.38–1.11)	1.05 (0.52–1.51)	1.33 (0.66–2.09)	1.50 (0.78–2.34)
50% clover	0.37 (0.22–0.78)	0.72 (0.36–1.11)	1.07 (0.51–1.50)	1.36 (0.65–2.10)	1.57 (0.79–2.39)
75% clover	0.41 (0.25–0.82)	0.78 (0.38–1.15)	1.16 (0.52–1.56)	1.49 (0.67–2.19)	1.73 (0.81–2.77)
100% clover	0.41 (0.25–0.84)	0.84 (0.38–1.20)	1.29 (0.54–1.83)	1.68 (0.70–2.75)	1.99 (0.85–3.57)
Empirically estimated N$_2$O emission[a] (kg N$_2$O-N ha^{-1} yr^{-1})					
IPCC	0.0 (0.0–0.0)	1.5 (0.5–4.5)	3.0 (0.9–9.0)	4.5 (1.4–13.5)	6.0 (1.8–18.0)
Shcherbak et al. (2014)[b]	0.0	2.0	5.5	10.6[c]	17.4[c]

Note. For comparison, we include the IPCC N$_2$O emission estimate with its 95% confidence intervals as well as the empirical estimates by Shcherbak et al. (2014). [a]For models empirically estimating N$_2$O emission, no distinction between different clover proportions is available. [b]Empirical model for "perennial grass/forage" in Table S3 of Shcherbak et al. (2014); emission of N$_2$O-N = 1.067 a (7.76 a 0.0353*Nfert)*Nfert) *10^{-3}. [c]Even though Shcherbak et al. (2014) displayed N$_2$O emission estimates for only N up to 300 kg N in their paper, the N fertilizer amounts in grasslands used for deriving their empirical relationship covered the higher levels (see supplementary information of Shcherbak et al., 2014), and we thus included levels up to 600 kg in our table.

3.2. Sward Productivity

Annual yields of the grass-clover mixtures ranged between 0.9 and 15.7 t DM ha^{-1} yr^{-1} across models and sites, clover, and fertilizer combinations (Figure 4a, Table 2). For APSIM, average yields across scenarios (i.e., variations in fertilizer inputs and clover proportion) ranged between 8.3 t DM ha^{-1} yr^{-1} at Site G1 and 11.9 t DM ha^{-1} yr^{-1} at Site G2 (Table 1). DayCent simulated lower average yields, ranging between 5.0 t DM ha^{-1} yr^{-1} (G1) and 8.8 t DM ha^{-1} yr^{-1} (G5) across scenarios (Table 1). Variability between scenarios (within model) was higher for APSIM (mean sum of squares MSQ$_{Nfert}$ = 890; MSQ$_{Clover\%}$ = 238.4) compared to DayCent (MSQ$_{Nfert}$ = 24.6, MSQ$_{Clover\%}$ = 24.6; Table S6). The models agreed that highest productivity levels were simulated at G2 and G5, medium productivity at G3, while lower productivity was simulated for G4 and G1 (Figure 4a, Table 1).
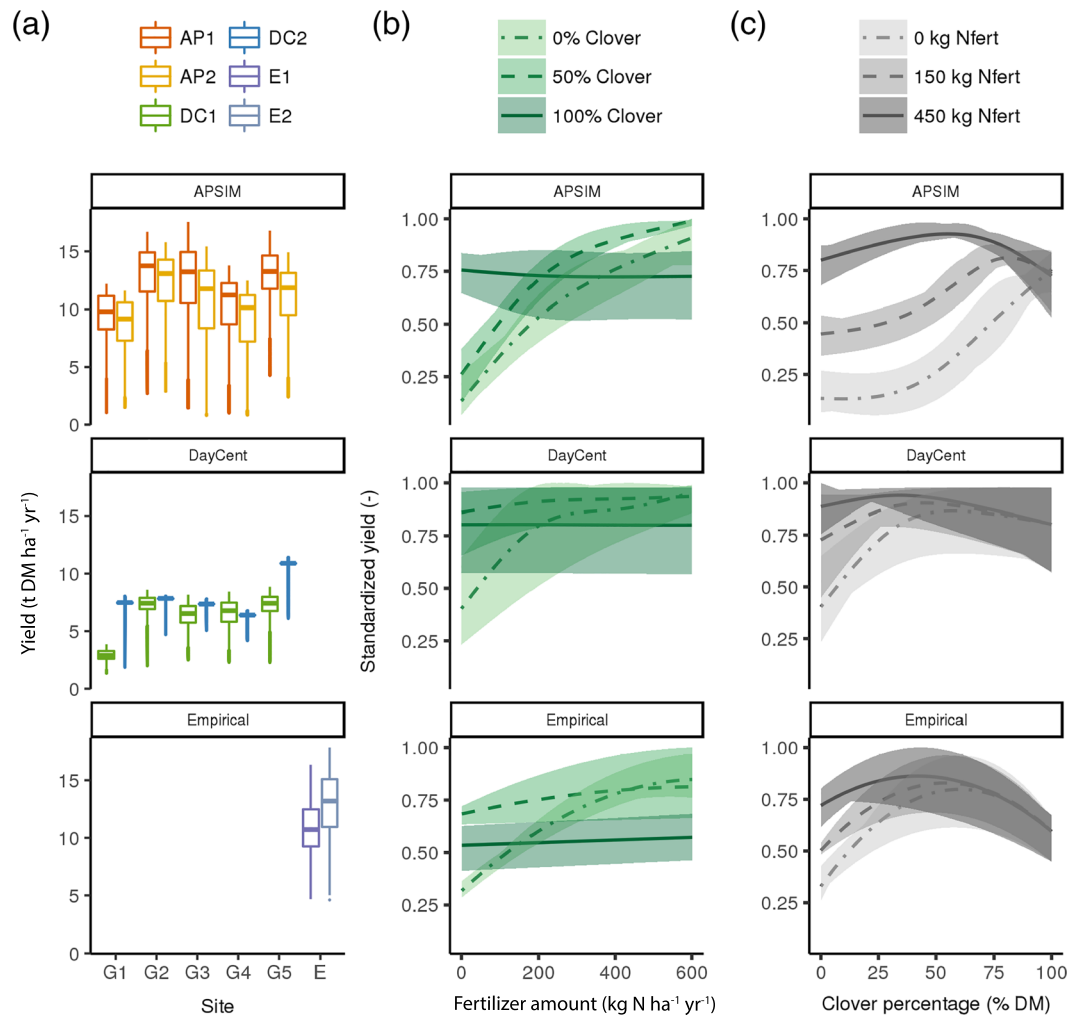
Figure 4. (a) Boxplots of yields across all mineral N scenarios (clover-fertilization combinations) per site (G1–G5) for APSIM (top) and DayCent variants (middle); (b) dependency of standardized yield (vertical axis) on fertilizer amount (horizontal axis) at 0% clover percentage (dot-dashed), 50% clover percentage (dashed), and 100% clover percentage (solid line); and (c) dependency of standardized yield (vertical axis) on clover percentage (horizontal axis) for the fertilizer levels 0 kg N (dot-dashed), 150 kg N (dashed), and 450 kg N (solid line) (APSIM: top, DayCent: middle). Yields in b and c were standardized with the maximum yields per model-site combination. Lines display the mean across sites, and shaded areas indicate the range across sites per model family. The bottom panels display conceptually our expectations from empirical findings. For the effect of fertilization (b), we displayed the empirical-derived regression equation from Reid (1983) (E1) and refer to both equation and coefficients in Table 2 in their publication. In contrast to our model outputs, grass-clover mixtures in Reid (1983) did not contain exactly 50% clover in dry biomass. For the effect of clover percentage (c), we used the empirical-derived equation from Nyfeler et al. (2009) (E2), equation 1 with the coefficients according to their Table S3 for mixtures of *T. repens* and *L. perenne*, excluding the other species. The Nyfeler study had 50 kg N as the minimum fertilizer rate, displayed dot-dashed, not 0 kg N. For the empirical studies in the bottom panel, the line indicates the mean and the shaded area the range across years.

3.2.1. Effects of Fertilizer Amount on Yield

Yields of grass monocultures significantly increased with the amount of fertilizer added. Average yields were lowest for scenarios without fertilizer (5.3 t DM ha^{-1} yr^{-1} in APSIM and 5.8 t DM ha^{-1} yr^{-1} in DayCent and up to 13.3 t DM ha^{-1} yr^{-1} in APSIM and 7.3 t DM ha^{-1} yr^{-1} in DayCent at 600 kg N input when averaged across sites and clover proportions; Table 1). While ASPIM suggested that increased fertilization still increased yields at the highest fertilizer levels (i.e., significant yield differences between 450 and 600 kg N$_{fert}$), DayCent simulated yields leveling off above ~200 kg N fertilization (i.e., no significant yield increase at the fertilizer levels ≥300 kg N compared to the 150 kg N; Table 1, Figure 4b). In unfertilized grass monocultures, APSIM simulated yields of 2.0 t DM ha^{-1} yr^{-1} (0.9–3.6 t DM ha^{-1} yr^{-1} across sites, Table 2)

Table 3
Average and Standard Deviation Across Sites (sd) of N_{symb} for Different Fertilizer Rates and Clover Percentages as Simulated by APSIM and DayCent

N_{fert} (kg ha^{-1} yr^{-1})	Clover percentage (%)	APSIM		DayCent	
		N_{symb} (kg ha^{-1} yr^{-1})	sd N_{symb} (kg ha^{-1} yr^{-1})	N_{symb} (kg ha^{-1} yr^{-1})	sd N_{symb} (kg ha^{-1} yr^{-1})
0	0	0	0	0	0
0	20	9	10	211	78
0	50	96	27	377	139
0	70	231	66	422	153
0	100	465	65	420	153
150	0	0	0	0	0
150	20	48	11	190	87
150	50	168	34	309	143
150	70	285	50	331	157
150	100	316	63	306	158
300	0	0	0	0	0
300	20	66	16	152	87
300	50	174	40	234	137
300	70	241	52	244	149
300	100	203	49	211	145

corresponding to 13% (7–21%) of the maximum site yields. In contrast, DayCent simulated higher yields of 3.2 t DM ha^{-1} yr^{-1} (1.6–4.2 t DM ha^{-1} yr^{-1}) or 36% (20–46%) of maximum site yields without any fertilization (reflected in the offset in Figure 4b).

The effect of fertilizer on yields depended on the clover proportion in the sward. In APSIM, a fertilizer response was found for grass-clover mixtures of 0–80% clover; for example, grass-clover mixtures of 50% in APSIM largely benefited from fertilization and increased yields from 3.9 t DM ha^{-1} yr^{-1} without fertilization to 12.2 t DM ha^{-1} yr^{-1} at 300 kg N fertilization, while high-clover mixtures ≥75% and clover monocultures did not benefit from fertilization above 300 kg N (Figure 4b, Table 2). DayCent showed a fertilizer response for mixtures of 25% clover, but little fertilization effect on mixtures of ≥50% clover (Figure 4b, Table 2), for example, for 50% clover, an increase in yield from 7.0 t DM ha^{-1} yr^{-1} without fertilization to 7.4 t DM ha^{-1} yr^{-1} at 300 kg N fertilization (Figure 4b, Table 2). Legume monocultures produced similar yields regardless of fertilizer level, consistently for all models (n.s. in Table 1), while the absolute amount of legume monoculture yields varied largely across models and sites. For example, APSIM predicted an average yield of 10.7 t DM ha^{-1} yr^{-1} ranging from 7.0–12.8 t DM ha^{-1} yr^{-1} across sites and DayCent 6.5 t DM ha^{-1} yr^{-1} with sites ranging from 4.9–8.5 t DM ha^{-1} yr^{-1} (Table 2).

3.2.2. Effect of Clover Proportion on Yields

Species mixtures were significantly more productive than grass monocultures in unfertilized swards (Table 2, Figure 4c). Transgressive over-yielding, that is, higher yields for mixtures than for the highest-yielding monoculture, were observed for DC1 and for APSIM (both variants) above 50 kg N fertilization. Without fertilization, maximum yields were found at 100% clover. When increasing fertilizer levels, the clover proportion for maximum production ("optimal clover proportion"—Clo_{opt}) shifted to lower clover proportions. For example, Clo_{opt} in APSIM was 82% (site-median) at 150 kg N, decreasing to 47% at 450 kg N, and in DayCent Clo_{opt} decreased from 58% at 150 kg N to 43% at 450 kg N.

3.3. Symbiotically Fixed Nitrogen

Symbiotically fixed nitrogen reached up to 465 (±65) kg N ha^{-1} yr^{-1} and 422 (±153) kg N ha^{-1} yr^{-1} at the scenarios without fertilizer inputs for APSIM and DayCent simulations, respectively (Table 3). With increasing clover proportion, N_{symb} increased up to a maximum and decreased at higher clover proportions. This pattern resulted from high productivity in mixtures due to the complementarity between grass and clover growth, which becomes less beneficial at higher clover proportions, resulting in smaller total growth and in less N_{symb}.

Consistent with the experimental findings of Nyfeler et al. (2009), the higher the amount of fertilizer, the lower the clover proportion at maximum N_{symb}. For example, in APSIM, the maximum N_{symb} at 150 kg N_{fert} was found at 86% clover (median across sites), while at fertilizer levels of 300 kg N, highest amounts of N_{symb} were found at 79% clover proportion. In DayCent, the maximum N_{symb} was found at slightly lower clover proportions around 70% clover for 150 kg N_{fert} and at 300 kg $N_{fert.}$

Fertilization increased N_{symb} for clover mixtures (below 70–80% clover) at moderate fertilizer levels as N addition stimulated growth. For example, grass-clover mixtures of 50% clover showed maximum fixation at around 150 kg N ha^{-1} yr^{-1} fertilization, while N_{symb} decreased with further fertilizer increment (Table 3). N_{symb} in clover monocultures decreased with increasing fertilization, as fertilized clover swards acquire their N from fertilization if available (Table 3).

3.4. Nitrous Oxide Emission Intensities

Nitrous oxide emission intensity estimates (N_2O-Int), that is, N_2O emissions per unit of biomass yield, allow a productivity-based comparison of N_2O emissions across sites, scenarios, and models. N_2O-Int ranged between 0.01 and 0.31 g N_2O-N kg^{-1} DM (APSIM) and 0.63–1.99 g N_2O-N kg^{-1} DM (DayCent) across clover proportions and fertilizer levels on average across sites (Table 2; Figure S3). Maximum N_2O-Int were reached at 100% clover and the maximal fertilizer rate (600 kg N_{fert}), except for DayCent at the G3 and G1 sites, where the maximum was already reached at 440 and 260 kg N_{fert} due to the leveling-off in N_2O emissions at higher fertilizer levels. Without fertilizer input, N_2O emission intensities maximized in grass monocultures which was primarily due to the small denominator (low yields) without fertilizer or N_{symb}. Significant variation in N_2O-Int among models, fertilizer inputs, and sites were observed. Models agreed that N_2O-Int were significantly lower for the sites G1 and G2 compared to G4 and G5 (Table 1). Increasing fertilizer inputs strongly increased N_2O emission intensity from 0.006 g N_2O-N kg^{-1} DM in unfertilized swards to 0.096 g N_2O-N kg^{-1} DM at 450 kg N fertilization for APSIM and from 0.341 g N_2O-N kg^{-1} DM at unfertilized sward to 1.232 g N_2O-N kg^{-1} DM at 450 kg N_{fert} for DayCent (Table 1, Figure S3). Lower N_2O-Int was found for mixtures compared to clover and grass monocultures (based on means across fertilizer levels, models, and sites, Figure S3, Table 1). Modeled N_2O-Int was highest when N input came from fertilizer and decreased with increasing fraction of the total N input acquired via symbiosis. While the pattern of lower N_2O intensities at higher fractions of N_{symb} was independent of the level of total N input, the magnitude (slope) of the effect on N_2O intensity was such that a great mitigation effect was achieved in systems with high total N input. Variability in N_2O-Int between models exceeded variability across sites and fertilizer levels meaning that there is considerable uncertainty in the estimated magnitude. Still, models agreed on the overall patterns of mixtures resulting in lower N_2O-Int compared to monocultures.

3.5. N_2O Mitigation Potential

In order to assess a mitigation potential, a baseline scenario is required. Here, we selected the scenario without clover and with a fertilizer rate of 200 kg N ha^{-1} yr^{-1} (N200.Clo0) as outlined above (see section 2.4). To understand the effect of successive reductions in fertilizer input, we assessed the results by two stepwise reductions of fertilizer to 100 and 0 kg N ha^{-1} yr^{-1} and determined the clover proportion that obtained the maximum yield (Clo$_{opt}$) for each site-model combination at the reduced fertilizer input. These were labeled Nx. Clo$_{opt}$ where x is the amount of fertilizer input.

Simulations suggested that N200.Clo0 soils emitted 0.46 ($\pm$0.27 SE) g N_2O-N per kg of dry matter yield across all models and sites (Figure 5). For the first reduction in fertilizer input, N100.Clo$_{opt}$, models showed consistently lower N_2O intensities of 0.30 ($\pm$0.25) g N_2O-N per kg of dry matter yield, corresponding to an average reduction of N_2O-Int by 34%. The lower N_2O-Int was present at all model-site combinations and resulted from a reduction of N_2O emissions by 25% on average and 32% higher yields compared to the baseline (Figure 5). Even though all model results showed that N_2O-Int consistently declined as fertilizer inputs were reduced, the magnitude of these reductions varied strongly across models and considerably across sites (Figure 5). A complete fertilizer omission reduced N_2O-Int to 0.19 g N_2O-N per kg of dry matter yield, corresponding to a 58% reduction compared to the N200.Clo0 baseline.

In order to define a target for the mitigation management, a key question is how clover proportions need to be adapted in order to produce maximum yields. Clover proportions at which the maximum yields were
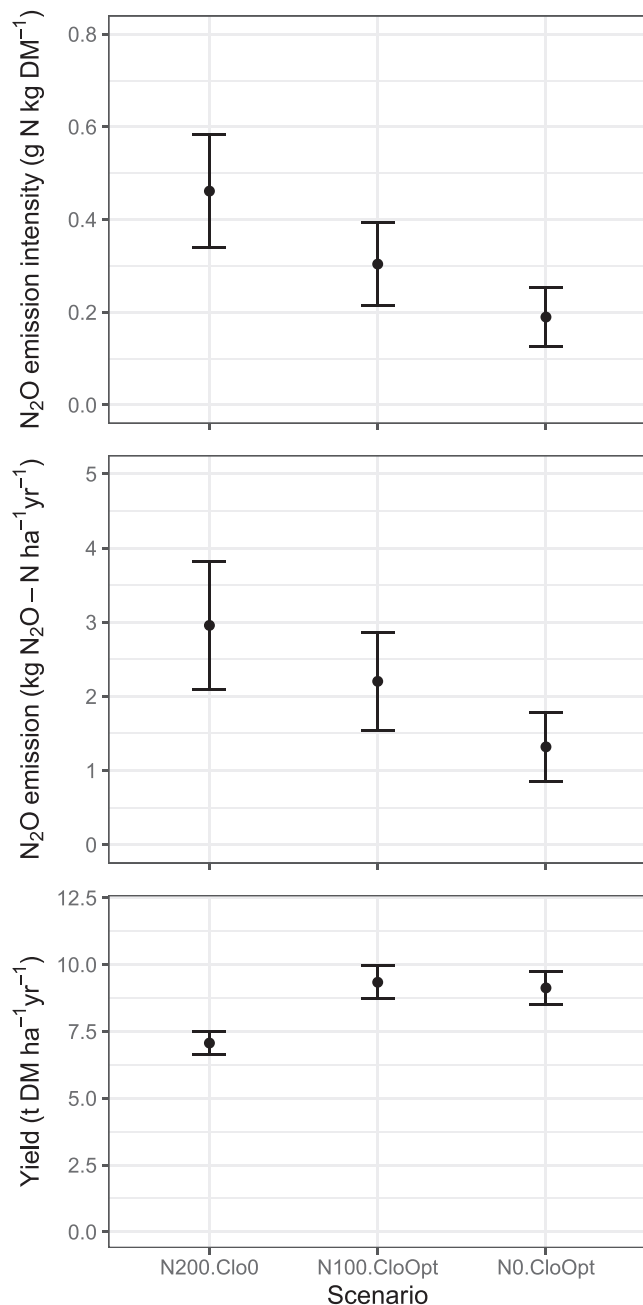
Figure 5. N_2O mitigation scenarios for mineral fertilizer replacement: (a) N_2O emission intensity; (b) N_2O emission; and (c) yields are displayed for scenario N200.Clo0 (200 kg N, 0% clover), N100.Clo$_{opt}$ (100 kg N, clover % for maximum yield), and N0.Clo$_{opt}$ (no fertilizer, clover % for maximum yield); the points display the mean, and whiskers display the standard error of the mean, calculated from the variability across models and sites.

reached ranged 66–86% (APSIM range across sites) and 0–72% (DayCent) for a N200.Clo$_{opt}$ scenario. A decrease in fertilization, N100.Clo$_{opt}$, gave higher optimal clover proportions (72–100% by APSIM and 40–100% by DayCent across sites) as might be expected. Complete fertilizer omission (N0.Clo$_{opt}$) resulted in clover proportions for optimal yield of 88–100% (APSIM) and 46–89% (DayCent).

If a reduction of N_2O emissions is accompanied by yield losses, its acceptability by farmers is likely to be very low. Thus, our aim was to identify scenarios at which the yield was not compromised by the N_2O mitigation and distinguish them from scenarios with yield losses. Surprisingly, yield reductions compared with N200.Clo0 were only found for N0.Clo$_{opt}$ at five out of 20 model-site combinations. Thus, these results suggest that a reduction of fertilizer from 200 to 0 kg N was possible without yield losses provided it was also possible to obtain an optimal clover proportion in the sward.

The global N_2O mitigation potential of replacing N_{fert} partially with N_{symb} was calculated for the optimal clover content and the 30% and 60% clover scenarios. Average N_2O emissions across our models were 2.96 kg N_2O-N ha^{-1} yr^{-1} if business-as-usual fertilizer application was 200 kg N ha^{-1} yr^{-1}. If the relative mitigation effects assessed our study are assumed to be representative for C3 pasturelands and assuming the baseline fertilizer application of 200 kg N ha^{-1} yr^{-1}, a reduction to 100 N ha^{-1} yr^{-1} combined with an increase to 30% or 60% clover would reduce fertilizer N_2O emissions by 37–40% and would scale to a global mitigation effect of 119–128 Gg N_2O-N yr^{-1}. While N_2O emissions were similar at 30% and 60% clover, yields were reduced (by 20%) at the 30% clover scenario compared to the baseline, while no yield decrease was found for the 60% clover scenario.

In comparison, a complete omission of N_{fert} combined with an increase in N_{symb} to its maximum scales to reductions of 258 Gg N ha^{-1} yr^{-1} for C3 pasturelands globally, when assuming 200 kg N yr^{-1} initial fertilization, respectively. Using initial clover percentage of 5%, instead of 0%, did not substantially affect these results.

4. Discussion

4.1. Yields, N_2O Emissions, N_2O Emission Intensities, and Mitigation Potentials

The general pattern of modeled yields increasing with increasing fertilizer input is consistent with experimental data. Yields of grass monocultures and grass-legume mixtures in APSIM increased across the whole range of fertilizer levels from 0 to 600 kg N ha^{-1} yr^{-1}; however, in DayCent, yields increased only between 0 to 200 kg N ha^{-1} yr^{-1} of fertilizer. Empirical findings show yields leveling off at high fertilizer amounts (400–500 kg N ha^{-1} yr^{-1}), depending on the site conditions (Prins, 1983; Reid, 1983; Sparrow, 1979; Whitehead, 2000) or still increasing at the highest N fertilizer levels (Wilkins et al., 2001). We conclude that yield responses to fertilization were more realistically simulated in APSIM but were unrealistically low in DayCent, which showed only marginally increased yield with increased fertilizer applications above 200 kg N ha^{-1} yr^{-1} at 50% clover (Table 2 even at the sites with high yield potential (G2–G5) (Prins, 1983; Whitehead, 2000). Underestimating a fertilizer effect on yields has a consequence for calculating mitigation potentials (section 3.5), as this would lead to an overestimation of potential reductions in N_2O-Int.

Model outputs reflected the influence of clover proportion on yields observed in grass-clover studies (Kirwan et al., 2007; Lüscher et al., 2014; Nyfeler et al., 2009; Suter et al., 2015). Yield maxima for DayCent were observed when the proportion of clover was between 38% and 79% across sites at 150 kg N and at relatively low clover proportions (0–47%) at 300 kg N. APSIM simulated yield maxima at comparably high clover proportions of 74–87% at 150 kg N and 66–76% at 300 kg N. Measured data showed overyielding at 40–60% clover for fertilizer levels of 50–450 kg ha^{-1} yr^{-1} at a site near G5 (Nyfeler et al., 2009), indicating that models were generally in line with observed data, but APSIM had a tendency to simulate higher optimal clover proportions than observed at this site. APSIM and DayCent differed with respect to plant N uptake at higher fertilizer levels. APSIM simulated highly efficient N uptake at high fertilizer levels, whereas in DayCent grasslands did not profit from increased fertilization above 200 kg N, and DayCent simulated increased leaching. It should be noted that the effect of legume percentage on DayCent's potential N fixation was derived from APSIM simulations (Fitton et al., 2019) while DayCent reduced N fixation from the potential based on its internal simulation of soil nitrogen conditions. Thus, while the two models used here are not completely independent, there are several degrees of separation between them. The development of a more process-based representation of multiple species in DayCent could avoid such a dependency.

Modeled N$_{symb}$ was within realistic ranges across all sites and scenarios (Table 1; Ledgard & Steele, 1992; Lüscher et al., 2014; Nyfeler et al., 2011), for example, N$_{symb}$ of up to 323 kg N ha^{-1} yr^{-1} was measured in Nyfeler et al. (2011) and values of up to 545 kg N ha^{-1} yr^{-1} have been observed by Carlsson and Huss-Danell (2003), both with *T. repens*, in line with our findings of maxima ranging between 424 and 559 kg N ha^{-1} yr^{-1} across sites. Grass-clover mixtures of <80% clover have the potential to fix more N$_2$ from the atmosphere compared to clover monocultures due to niche complementarity, while clover proportions >80% are less beneficial due to less growth and therefore less N$_{symb}$ (Nyfeler et al., 2011), which was also depicted by the models.

Simulated N$_2$O emissions in APSIM were low compared to N$_2$O emission estimates using IPCC Emission Factors at all fertilizer levels (i.e., below the IPCC lower confidence interval). These low N$_2$O emissions could arise either from APSIM simulating low soil organic matter turnover rates or because APSIM simulates N being taken up by plants too efficiently or a combination of both effects. In contrast, N$_2$O emissions in DayCent exceeded IPCC mean estimates at all fertilizer levels but were within the upper confidence interval estimates of the IPCC and were similar to estimates from Shcherbak et al. (2014) for the upper fertilizer application rates (>300 kg N ha^{-1} yr^{-1}). These high N$_2$O emissions in DayCent coincided with increased leaching rates at higher levels of fertilization, reflecting that plants were not able to benefit from additional N fertilization and thus biomass yields from DayCent at these fertilizer rates seemed to be unrealistically low. While the current IPCC Emission Factor (EF1) approach assumes a linear increase in N$_2$O emissions with N fertilizer amounts, Shcherbak et al. (2014) found N$_{fert}$ to have a nonlinear effect on N$_2$O emissions. They analyzed site data across the globe using multiple N application rates, showing that the increases in N$_2$O emissions were higher at higher fertilizer amounts (mean +0.0033% change in N$_2$O emitted per additional kg N ha^{-1} yr^{-1} of fertilizer added based on 41 site-years for grassland; Table S3 in Shcherbak et al., 2014). Greater change in N$_2$O emissions per N$_{fert}$ increment at higher fertilizer input levels implies that N$_2$O emission reductions are more effective in high N input systems compared to systems of moderate or low fertilizer N inputs, where a reduction in the same amount of N$_{fert}$ is less effective. Kim et al. (2013) conceptualized the relationship between N fertilizer amounts and N$_2$O emissions and hypothesized an initial linear increase of N$_2$O production at the low fertilizer levels at which microbes compete with plants for available N. Additional N fertilizer increase beyond optimal plant N uptake rates would then lead to an exponential increase in N$_2$O production due to an excess of available N, stimulating nitrification and denitrification. Eventually, further fertilizer N increment would lead to N$_2$O production leveling off, leading to a steady state where soil C limits N$_2$O production (Kim et al., 2013; Rochette et al., 2010). Model outputs showed nonlinear (APSIM) and linear (DayCent) effects of N fertilizer amounts on N$_2$O emissions on average, while leveling off was observed for some model-site combinations in DayCent.

In line with the findings reviewed in Lüscher et al. (2014), N$_2$O-Int was lower in grass-clover mixtures compared to monocultures, which was expected since N$_2$O-Int penalizes yield losses (as observed at low N input), and penalizes high N$_2$O emissions (as observed at high N inputs). Thus, grass-clover mixtures optimize the combination of both. As the increase in N$_2$O emissions with increases in N fertilizer typically

exceed the yield gain for the same fertilizer increment, N_2O-Int were minimized at the lowest fertilizer levels (Table 1). The modeled results of yields being maintained under reduced fertilizer might be too optimistic. However, in a recent field experiment at Site G5 (Fuchs et al., 2018), N_2O-Int was reduced from 0.42 g N kg DM^{-1} yr^{-1} on a fertilized control obtaining fertilizer N via slurry (240 kg N, 4–15% clover) to 0.28 g N kg DM^{-1} yr^{-1} at the clover treatment (23–40% clover and no N_{fert}) over a 2-year observation period. Further, lab experiments by Barneze et al. (2019) confirmed the model results that higher clover proportions increase productivity without affecting N_2O emissions, resulting in reduced N_2O-Int.

The yields at higher clover percentages (75%) were overestimated by DayCent and APSIM (Table S1b). While DayCent predicted an optimal clover content around 50%, similar to the observations, APSIM predicted higher optimum clover content around 75%. The optimum clover proportion thus could likely be overestimated in this modeling exercise.

Dangal et al. (2019) modeled total emissions from grassland of 3.11 Tg N_2O-N; other estimates range from 1.86 Tg N_2O-N in the EDGAR database (EDGAR, 2018) to 2.8 Tg N_2O-N in the study of Davidson (2009, referring to 2000). A mitigation potential of 128 Gg N_2O-N yr^{-1} (which resulted from scaling up the results from our study; baseline 200 kg N) resulted in a 4.1% reduction of global grassland N_2O emissions if the Dangal et al. estimate is used a reference. The tested mitigation option reduced total grassland N_2O emissions only moderately, since the largest fraction of grassland N_2O emissions is attributed to the fate of animal excreta and manure (Dangal et al., 2019), while fertilizer application contributes 11%. The estimated average N_2O emissions for C3 pasture in Dangal et al. (2019) was rather high with an average across C3 pasturelands of 3.95 kg N_2O-N ha^{-1} yr^{-1} during 2010–2014 and was comparable to the estimates by our models for grasslands receiving 300 kg N_{fert} ha^{-1} yr^{-1}.

Replacing not only synthetic fertilizers but also organic N application and applying the organic N on croplands would at least double the estimated N_2O mitigation effect, since higher emissions come from organic fertilizers compared to mineral fertilizers (Dangal et al., 2019). Modeling the scenarios presented here with organic fertilizer instead of mineral fertilizer resulted in similar reductions of N_2O. If the surplus of organic fertilizer could replace mineral fertilizer on croplands, the mitigation strategy could have far more impact. Since excreta deposition contributes the large fractions (54%; Dangal et al., 2019) of grassland N_2O emissions, larger reductions of grassland N_2O emissions need to target excreta of grazing livestock.

The global estimates of N_2O emission reductions largely depend on the absolute N_2O emissions estimated by the applied models. We are more confident in the relative estimates compared to absolute estimates, as the mitigation potentials largely are driven by the N_2O emission reductions due to omitted N fertilization. Besides the model uncertainties, the global estimates contain uncertainty, first due to the strong dependence on the outcomes of Dangal et al. (2019) and second due to the use of an average mitigation potential for a limited number ($n = 5$) of sites. Thus, these estimates should rather be seen as an indicator for the magnitude, keeping these uncertainties in mind.

4.2. Conclusion and Outlook

In this study, we investigated the effect of reduced fertilizer combined with increased clover proportions as a mitigation strategy and contributed to a mechanistic understanding of the relationships between the proportion of clover and N_2O emissions and emission intensities. Results from different models gave contrasting perspectives on N_2O emissions from grassland systems, with quite diverse outcomes in absolute N_2O emission intensities, but agreement on the overall pattern of N_2O emission reductions on a per-unit-area, as well as on a per-unit-product basis. Models predicted relative changes in N_2O emission and N_2O-Int (relative to the overall site-model reference level) much more consistently than they predicted the absolute changes in site-specific N_2O emission intensities. Consequently, a more robust statement on the relative reduction in N_2O emission intensity than on absolute N_2O reduction potential is possible.

The simulated effects of clover percentage and fertilizer on N_2O emissions agreed well with the empirical studies, that is, no increase in N_2O emissions was observed at low to medium fertilizer amounts with increasing clover percentage up to 60% (compare Figure 1b with Figure 3). Assuming 0–5% clover and 200 kg N ha^{-1} yr^{-1} on average as the baseline scenario, an increase in clover percentage of all fertilized

C3 pastures to 30% combined with a reduction of fertilizer inputs to 100 N ha^{-1} yr^{-1} would reduce global N$_2$O emissions by 128 Gg yr^{-1}, around 4% of global grassland N$_2$O emissions.

Different biogeochemical models tend to have complementary strengths, leading to their multi-model ensemble median performing better compared to individual models (Asseng et al., 2013; Ehrhardt et al., 2018). Thus, using an ensemble reduces the overall error of the predictions, while also providing an estimate of the uncertainty of those predictions. However, constructing a large ensemble was impossible for this exercise because only few models are capable of simulating species mixtures (Kipling et al., 2016; Snow et al., 2014). Major challenges were the adequate simulation of BNF in grass-clover-mixtures, which finally led to a reduced number of participating models. In order to tackle present shortcomings, models need to be improved and species-interactions need to be considered in order to better and more realistically reflect the ongoing processes. Few models adequately simulate grazing effects (Snow et al., 2014). We simulated cut-and-carry systems in order to tackle the complexity of the modeling and ensure interpretable outputs. In reality, grazing would impact grassland productivity, legume content, and N$_2$O emissions. However, the reasoning was to first consider the simpler system in order to enable understanding without confounding effects. An important next step for this work is to test the mitigation under grazing; however, this would entail first overcoming challenges with models and their abilities to reasonably simulate grazed conditions (e.g., see Snow et al., 2014).

Improved land use mapping of grass-legume mixtures applying machine learning algorithms on remotely sensed high-resolution data may in the future allow the extent of grass-legume mixtures to be identified, and at some point, also the respective percentage of legumes in the mixtures. Such detailed maps will enable a specific baseline on the actual extent and distribution of grass-legume mixtures to be defined and consequently enable a better estimate on the N$_2$O mitigation potential of grass-legume mixtures.

Practical limitations potentially restricting the applicability of the investigated N$_2$O mitigation strategy were not considered here. These limitations include the challenge of achieving high and persistent legume proportions, particularly under grazing. Further, clover is not well adapted for all locations, and local soil conditions such as low soil pH and waterlogging can severely limit clover production.

Here, we have shown that grass-clover mixtures with the lowest fertilizer input were associated with the lowest N$_2$O emission intensities. Experimental studies which systematically assess N$_2$O emissions under different clover proportions for different N fertilizer regimes would be highly beneficial. Experimental work should focus on assessing in detail swards with clover proportions of 30–50% and ≤150 kg N input as these seem most beneficial with respect to multiple outputs such as yields, N yields and also feeding values (Lüscher et al., 2014). Additional data on the endpoints (i.e., 0% and 100%) clover would assist in confirming or refuting the modeled outcomes. Legumes have been shown to perform better than non-legumes under drought conditions (Hofer et al., 2017) and show higher water-use efficiencies than non-legumes (Adams et al., 2018). Together with the fact that legumes can replace mineral fertilizer additions make them important species for increasing sustainability and resilience of agricultural production under changing climate, highlighting the role legumes could play not only for climate change mitigation but also for adaptation.

Data Availability Statement

The data used in this study are available in the ETH Research collection under https://www.research-collection.ethz.ch/handle/20.500.11850/395370 as "Dataset: Evaluating the potential of legumes to mitigate N$_2$O emissions from permanent grassland using process-based models."

References

Adams, M. A., Buchmann, N., Sprent, J., Buckley, T. N., & Turnbull, T. L. (2018). Crops, nitrogen and water: Are legumes friend, foe or misunderstood ally? *Trends in Plant Science*, *1655*, 1–12. https://doi.org/10.1016/j.tplants.2018.02.009

Ammann, C., Spirig, C., Leifeld, J., & Neftel, A. (2009). Assessment of the nitrogen and carbon budget of two managed temperate grassland fields. *Agriculture, Ecosystems & Environment*, *133*(3–4), 150–162. https://doi.org/10.1016/j.agee.2009.05.006

Asseng, S., Ewert, F., Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., et al. (2013). Uncertainty in simulating wheat yields under climate change. *Nature Climate Change*, *3*(9), 827–832. https://doi.org/10.1038/nclimate1916

Baggs, E. M. (2011). Soil microbial sources of nitrous oxide: Recent advances in knowledge, emerging challenges and future direction. *Current Opinion in Environmental Sustainability*, *3*(5), 321–327. https://doi.org/10.1016/j.cosust.2011.08.011

Barneze, A. S., Whitaker, J., McNamara, N. P., & Ostle, N. J. (2019). Legumes increase grassland productivity with no effect on nitrous oxide emissions. *Plant and Soil*, *446*. https://doi.org/10.1007/s11104-019-04338-w

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

Bolan, N. S., Saggar, S., Luo, J., Bhandral, R., & Singh, J. (2004). In D. Sparks (Ed.), *Gaseous emissions of nitrogen from grazed pastures: Processes, measurements and modelling, environmental implications, and mitigation* (1st ed., Vol. 84, pp. 37–120). San Diego, USA: Elsevier. https://doi.org/10.1016/S0065-2113(04)84002-1

Brentrup, F., & Pallière, C. (2008). GHG emissions and energy efficiency in European nitrogen fertiliser production and use. *Proceedings - International Fertiliser Society*, *639*, 1–25.

Cahalan, E., Ernfors, M., Müller, C., Devaney, D., Laughlin, R. J., Watson, C. J., et al. (2015). The effect of the nitrification inhibitor dicyandiamide (DCD) on nitrous oxide and methane emissions after cattle slurry application to Irish grassland. *Agriculture, Ecosystems & Environment*, *199*, 339–349. https://doi.org/10.1016/j.agee.2014.09.008

Carlsson, G., & Huss-Danell, K. (2003). Nitrogen fixation in perennial forage legumes in the field. *Plant and Soil*, *253*(2), 353–372. https://doi.org/10.1023/a:1024847017371

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., et al. (2013). Carbon and other biogeochemical cycles. In T. F. Stocker et al. (Eds.), *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 465–570). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, *83*(403), 596–610. https://doi.org/10.1080/01621459.1988.10478639

Dangal, S. R. S., Tian, H., Xu, R., Chang, J., Canadell, J. G., Ciais, P., et al. (2019). Global nitrous oxide emissions from pasturelands and rangelands: Magnitude, spatiotemporal patterns, and attribution. *Global Biogeochemical Cycles*, *33*(2), 200–222. https://doi.org/10.1029/2018GB006091

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

Davidson, E. A. (2009). The contribution of manure and fertilizer nitrogen to atmospheric nitrous oxide since 1860. *Nature Geoscience*, *2*(9), 659–662. https://doi.org/10.1038/ngeo608

de Klein, C. A. M., & Eckard, R. J. (2008). Targeted technologies for nitrous oxide abatement from animal agriculture. *Australian Journal of Experimental Agriculture*, *48*(1–2), 14–20. https://doi.org/10.1071/EA07217

Del Grosso, S. J., Parton, W. J., Moiser, A. R., Hartman, M. D., Brenner, J., Ojima, D. S., & Schimel, D. S. (2001). Simulated interaction of carbon dynamics and nitrogen trace gas fluxes using the DAYCENT mode. In M. J. Shaffer, L. Ma, & S. Hansen (Eds.), *Modeling carbon and nitrogen dynamics for soil management* (pp. 303–332). Boca Raton, FL: CRC Press. https://doi.org/10.1201/9780367801373/

Dowle, M., & Srinivasan, A. (2017). *data.table: Extension of 'data.frame'*. Retrieved from https://CRAN.R-project.org/package=data.table

Eckard, R. J., Grainger, C., & de Klein, C. A. M. (2010). Options for the abatement of methane and nitrous oxide from ruminant production: A review. *Livestock Science*, *130*(1), 47–56. https://doi.org/10.1016/j.livsci.2010.02.010

EDGAR (2018). *Emission database for global atmospheric research (EDGAR v4.3.2), release version 4.2*. Retrieved from http://edgar.jrc.ec.europa.eu

Ehrhardt, F., Soussana, J.-F., Bellocchi, G., Grace, P., McAuliffe, R., Recous, S., et al. (2018). Assessing uncertainties in crop and pasture ensemble model simulations of productivity and N₂O emissions. *Global Change Biology*, *24*. https://doi.org/10.1111/gcb.13965

Finn, J. A., Kirwan, L., Connolly, J., Sebastià, M. T., Helgadottir, A., Baadshaug, O. H., et al. (2013). Ecosystem function enhanced by combining four functional types of plant species in intensively managed grassland mixtures: A 3-year continental-scale field experiment. *Journal of Applied Ecology*, *50*(2), 365–375. https://doi.org/10.1111/1365-2664.12041

Fitton, N., Bindi, M., Newton, P., Smith, P., Lüscher, A., Huguenin-Elie, O., et al. (2019). Modelling biological N fixation and grass-legume dynamics with process-based biogeochemical models of varying complexity. *European Journal of Agronomy*, *642*. https://doi.org/10.1016/j.eja.2019.03.008

Flechard, C. R., Ambus, P., Skiba, U., Rees, R. M., Hensen, A., van Amstel, A., et al. (2007). Effects of climate and management intensity on nitrous oxide emissions in grassland systems across Europe. *Agriculture Ecosystems & Environment*, *121*(1–2), 135–152. https://doi.org/10.1016/j.agee.2006.12.024

Flessa, H. (2012). *Studie zur Vorbereitung einer effizienten und gut abgestimmten Klimaschutzpolitik für den Agrarsektor* (Vol. 361, pp. 1–472). Braunschweig, Germany: Johann Heinrich von Thünen-Institut. Retrieved from https://literatur.thuenen.de/digbib_extern/dn050716.pdf

Flessa, H., Greef, J. M., Hofmeier, M., Dittert, K., Ruser, R., Osterburg, B., et al. (2014). *Minderung von Stickstoff-Emissionen in der Landwirtschaft - Empfehlungen für die Praxis und aktuelle Fragen an die Wissenschaft* (Vol. 1). Berlin, Germany: Senat der Bundesforschungsinstitute des Bundesministeriums für Ernährung und Landwirtschaft. Retrieved from http://literatur.ti.bund.de/digbib_extern/dn054531.pdf

Fuchs, K., Hörtnagl, L., Buchmann, N., Eugster, W., Snow, V., & Merbold, L. (2018). Management matters: Testing a mitigation strategy for nitrous oxide emissions on intensively managed grassland. *Biogeosciences*, *15*, 5519–5543. https://doi.org/10.5194/bg-15-5519-2018

Fuchs, K., Merbold, L., Buchmann, N., Bretscher, D., Brilli, L., Fitton, N., et al. (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

Galbally, I. E., Meyer, M. C. P., Wang, Y.-P., Smith, C. J., & Weeks, I. A. (2010). Nitrous oxide emissions from a legume pasture and the influences of liming and urine addition. *Agriculture Ecosystems & Environment*, *136*(3–4), 262–272. https://doi.org/10.1016/j.agee.2009.10.013

Hofer, D., Suter, M., Buchmann, N., & Lüscher, A. (2017). Nitrogen status of functionally different forage species explains resistance to severe drought and post-drought overcompensation. *Agriculture Ecosystems & Environment*, *236*, 312–322. https://doi.org/10.1016/j.agee.2016.11.022

Holzworth, D. P., Huth, N. I., deVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., et al. (2014). APSIM—Evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software*, *62*, 327–350. https://doi.org/10.1016/j.envsoft.2014.07.009

Hörtnagl, L., Barthel, M., Buchmann, N., Eugster, W., Butterbach-Bahl, K., Díaz-Pinés, E., et al. (2018). Greenhouse gas fluxes over managed grasslands in Central Europe. *Global Change Biology*, *24*. https://doi.org/10.1111/gcb.14079

Huth, N. I., Bristow, K. L., & Verburg, K. (2012). SWIM3: Model use, calibration, and validation. *Transactions of the ASABE*, *55*(4), 1303–1313. https://doi.org/10.13031/2013.42243

Ingram, L. J., Schuman, G. E., Parkin, T. B., & Mortenson, M. (2015). Trace gas fluxes from a Northern mixed-grass prairie interseeded with Alfalfa. *Plant and Soil*, *386*(1-2), 285–301. https://doi.org/10.1007/s11104-014-2264-9

Jensen, E. S., Peoples, M. B., Boddey, R. M., Gresshoff, P. M., Hauggaard-Nielsen, H., Alves, B. J. R., & Morrison, M. J. (2012). Legumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries: A review. *Agronomy for Sustainable Development*, *32*(2), 329–364. https://doi.org/10.1007/s13593-011-0056-7

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

Kipling, R. P., Bannink, A., Bellocchi, G., Dalgaard, T., Fox, N. J., Hutchings, N. J., et al. (2016). Modeling European ruminant production systems: Facing the challenges of climate change. *Agricultural Systems*, *147*, 24–37. https://doi.org/10.1016/j.agsy.2016.05.007

Kirwan, L., Luescher, A., Sebastia, M. T., Finn, J. A., Collins, R. P., Porqueddu, C., et al. (2007). Evenness drives consistent diversity effects in intensive grassland systems across 28 European sites. *Journal of Ecology*, *95*(3), 530–539. https://doi.org/10.1111/j.1365-2745.2007.01225.x

Klein Goldewijk, K., Beusen, A., Doelman, J., & Stehfest, E. (2017). Anthropogenic land use estimates for the Holocene—HYDE 3.2. *Earth System Science Data*, *9*(2), 927–953. https://doi.org/10.5194/essd-9-927-2017

Klumpp, K., Bloor, J. M. G., Ambus, P., & Soussana, J.-F. (2011). Effects of clover density on N_2O emissions and plant-soil N transfers in a fertilised upland pasture. *Plant and Soil*, *343*(1–2), 97–107. https://doi.org/10.1007/s11104-010-0526-8

Lam, S. K., Suter, H., Mosier, A. R., & Chen, D. (2017). Using nitrification inhibitors to mitigate agricultural N_2O emission: A double-edged sword? *Global Change Biology*, *23*(2), 485–489. https://doi.org/10.1111/gcb.13338

Ledgard, S. F., & Steele, K. W. (1992). Biological nitrogen fixation in mixed legume/grass pastures. *Plant and Soil*, *141*(1–2), 137–153. https://doi.org/10.1007/BF00011314

Li, D., Lanigan, G., & Humphreys, J. (2011). Measured and simulated nitrous oxide emissions from ryegrass- and ryegrass/white clover-based grasslands in a moist temperate climate. *PLoS ONE*, *6*(10), e26176. https://doi.org/10.1371/journal.pone.0026176

Li, D., Watson, C. J., Yan, M. J., Lalor, S., Rafique, R., Hyde, B., et al. (2013). A review of nitrous oxide mitigation by farm nitrogen management in temperate grassland-based agriculture. *Journal of Environmental Management*, *128*, 893–903. https://doi.org/10.1016/j.jenvman.2013.06.026

Li, F., Snow, V., & Holzworth, D. (2011). Modelling the seasonal and geographical pattern of pasture production in New Zealand. *New Zealand Journal of Agricultural Research*, *54*(4), 331–352. https://doi.org/10.1080/00288233.2011.613403

Lugato, E., Leip, A., & Jones, A. (2018). Mitigation potential of soil carbon management overestimated by neglecting N_2O emissions. *Nature Climate Change*, *8*(3), 219–223. https://doi.org/10.1038/s41558-018-0087-z

Luo, J., de Klein, C. A. M., Ledgard, S. F., & Saggar, S. (2010). Management options to reduce nitrous oxide emissions from intensively grazed pastures: A review. *Agriculture, Ecosystems & Environment*, *136*(3–4), 282–291. https://doi.org/10.1016/j.agee.2009.12.003

Lüscher, A., Mueller-Harvey, I., Soussana, J. F., Rees, R. M., & Peyraud, J. L. (2014). Potential of legume-based grassland–livestock systems in Europe: A review. *Grass and Forage Science*, *69*(2), 206–228. https://doi.org/10.1111/gfs.12124

Mosier, A., Wassmann, R., Verchot, L., King, J., & Palm, C. (2004). Methane and nitrogen oxide fluxes in tropical agricultural soils: Sources, sinks and mechanisms. *Environment, Development and Sustainability*, *6*(1–2), 11–49. https://doi.org/10.1023/B:ENVI.0000003627.43162.ae

Niklaus, P. A., Le Roux, X., Poly, F., Buchmann, N., Scherer-Lorenzen, M., Weigelt, A., & Barnard, R. L. (2016). Plant species diversity affects soil-atmosphere fluxes of methane and nitrous oxide. *Oecologia*, *181*(3), 919–930. https://doi.org/10.1007/s00442-016-3611-8

Nyfeler, D., Huguenin-Elie, O., Suter, M., Frossard, E., Connolly, J., & Lüscher, A. (2009). Strong mixture effects among four species in fertilized agricultural grassland led to persistent and consistent transgressive overyielding. *Journal of Applied Ecology*, *46*(3), 683–691. https://doi.org/10.1111/j.1365-2664.2009.01653.x

Nyfeler, D., Huguenin-Elie, O., Suter, M., Frossard, E., & Lüscher, A. (2011). Grass–legume mixtures can yield more nitrogen than legume pure stands due to mutual stimulation of nitrogen uptake from symbiotic and non-symbiotic sources. *Agriculture, Ecosystems & Environment*, *140*(1–2), 155–163. https://doi.org/10.1016/j.agee.2010.11.022

Osterburg, B., Rüter, S., Freibauer, A., de Witte, T., Elsasser, P., Kätsch, S., et al. (2013). *Handlungsoptionen für den Klimaschutz in der deutschen Agrar-und Forstwirtschaft*. Braunschweig: Thünen Report. Retrieved from http://www.econstor.eu/handle/10419/88588

Parton, W. J., Holland, E. A., Grosso, S. J. D., Hartman, M. D., Martin, R. E., Mosier, A. R., et al. (2001). Generalized model for NO_x and N_2O emissions from soils. *Journal of Geophysical Research*, *106*(D15), 17,403–17,419. https://doi.org/10.1029/2001JD900101

Parton, W. J., Hartman, M., Ojima, D., & Schimel, D. (1998). DAYCENT and its land surface submodel: Description and testing. *Global and Planetary Change*, *19*(1–4), 35–48. https://doi.org/10.1016/S0921-8181(98)00040-X

Phelan, P., Moloney, A. P., McGeough, E. J., Humphreys, J., Bertilsson, J., O'Riordan, E. G., & O'Kiely, P. (2015). Forage legumes for grazing and conserving in ruminant production systems. *Critical Reviews in Plant Sciences*, *34*(1–3), 281–326. https://doi.org/10.1080/07352689.2014.898455

Prins, W. H. (1983). *Limits to nitrogen fertilizer on grassland (PhD thesis)*. Wageningen: Wageningen University. Retrieved from http://library.wur.nl/WebQuery/wurpubs/77523

Probert, M. E., Dimes, J. P., Keating, B. A., Dalal, R. C., & Strong, W. M. (1998). APSIM's water and nitrogen modules and simulation of the dynamics of water and nitrogen in fallow systems. *Agricultural Systems*, *56*(1), 1–28. https://doi.org/10.1016/s0308-521x(97)00028-0

Ravishankara, A. R., Daniel, J. S., & Portmann, R. W. (2009). Nitrous oxide (N_2O): The dominant ozone-depleting substance emitted in the 21st century. *Science*, *326*(5949), 123–125. https://doi.org/10.1126/science.1176985

R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/

Reid, D. (1983). The combined use of fertilizer nitrogen and white clover as nitrogen sources for herbage growth. *The Journal of Agricultural Science*, *100*(3), 613–623. https://doi.org/10.1017/S0021859600035371

Rochette, P., Tremblay, N., Fallon, E., Angers, D. A., Chantigny, M. H., MacDonald, J. D., et al. (2010). N_2O emissions from an irrigated and non-irrigated organic soil in eastern Canada as influenced by N fertilizer addition. *European Journal of Soil Science*, *61*(2), 186–196. https://doi.org/10.1111/j.1365-2389.2009.01222.x

Rochette, P., & Janzen, H. H. (2005). Towards a revised coefficient for estimating N_2O emissions from legumes. *Nutrient Cycling in Agroecosystems*, *73*(2–3), 171–179. https://doi.org/10.1007/s10705-005-0357-9

Ruane, A. C., Goldberg, R., & Chryssanthacopoulos, J. (2015). Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. *Agricultural and Forest Meteorology*, *200*, 233–248. https://doi.org/10.1016/j.agrformet.2014.09.016

Saggar, S., Bolan, N. S., Bhandral, R., Hedley, C. B., & Luo, J. (2004). A review of emissions of methane, ammonia, and nitrous oxide from animal excreta deposition and farm effluent application in grazed pastures. *New Zealand Journal of Agricultural Research*, *47*(4), 513–544. https://doi.org/10.1080/00288233.2004.9513618

Saggar, S., Jha, N., Deslippe, J., Bolan, N. S., Luo, J., Giltrap, D. L., et al. (2013). Denitrification and $N_2O:N_2$ production in temperate grasslands: Processes, measurements, modelling and mitigating negative impacts. *Science of the Total Environment*, *465*, 173–195. https://doi.org/10.1016/j.scitotenv.2012.11.050

Sándor, R., Ehrhardt, F., Brilli, L., Carozzi, M., Recous, S., Smith, P., et al. (2018). The use of biogeochemical models to evaluate mitigation of greenhouse gas emissions from managed grasslands. *Science of the Total Environment*, *642*, 292–306. https://doi.org/10.1016/j.scitotenv.2018.06.020

Schmeer, M., Loges, R., Dittert, K., Senbayram, M., Horn, R., & Taube, F. (2014). Legume-based forage production systems reduce nitrous oxide emissions. *Soil and Tillage Research*, *143*, 17–25. https://doi.org/10.1016/j.still.2014.05.001

Shcherbak, I., Millar, N., & Robertson, G. P. (2014). Global metaanalysis of the nonlinear response of soil nitrous oxide (N_2O) emissions to fertilizer nitrogen. *Proceedings of the National Academy of Sciences*, *111*(25), 9199–9204. https://doi.org/10.1073/pnas.1322434111

Simek, M. (2004). Emissions of N_2O and CO_2, denitrification measurements and soil properties in red clover and ryegrass stands. *Soil Biology and Biochemistry*, *36*(1), 9–21. https://doi.org/10.1016/j.soilbio.2003.08.010

Smith, K. A. (2017). Changing views of nitrous oxide emissions from agricultural soil: Key controlling processes and assessment at different spatial scales: N_2O emissions from agricultural soil. *European Journal of Soil Science*, *68*(2), 137–155. https://doi.org/10.1111/ejss.12409

Smith, K. A., Dobbie, K. E., Thorman, R., Watson, C. J., Chadwick, D. R., Yamulki, S., & Ball, B. C. (2012). The effect of N fertilizer forms on nitrous oxide emissions from UK arable land and grassland. *Nutrient Cycling in Agroecosystems*, *93*(2), 127–149. https://doi.org/10.1007/s10705-012-9505-1

Smith, K. A., McTaggart, I. P., & Tsuruta, H. (1997). Emissions of N_2O and NO associated with nitrogen fertilization in intensive agriculture, and the potential for mitigation. *Soil Use and Management*, *13*(s4), 296–304. https://doi.org/10.1111/j.1475-2743.1997.tb00601.x

Smith, P., Bustamante, M., Ahammad, H., Clark, H., Dong, H., Elsiddig, E. A., et al. (2014). Agriculture, forestry and other land use (AFOLU). In O. Edenhofer, (Eds.), *Climate change 2014: Mitigation of climate change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Chap. 11, pp. 811–922). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H., Kumar, P., et al. (2008). Greenhouse gas mitigation in agriculture. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *363*(1492), 789–813. https://doi.org/10.1098/rstb.2007.2184

Snow, V. O., & Huth, N. I. (2004). *The APSIM–MICROMET module, 2004 (12848) (Internal Report No. 2004/12848)* (pp. 1–18). Auckland, New Zealand: HortResearch. https://www.apsim.info/wp-content/uploads/2019/09/Micromet.pdf

Snow, V., Rotz, C. A., Moore, A. D., Martin-Clouaire, R., Johnson, I. R., Hutchings, N. J., & Eckard, R. J. (2014). The challenges—and some solutions—to process-based modelling of grazed agricultural systems. *Environmental Modelling & Software*, *62*, 420–436. https://doi.org/10.1016/j.envsoft.2014.03.009

Sparrow, P. E. (1979). The comparison of five response curves for representing the relationship between the annual dry-matter yield of grass herbage and fertilizer nitrogen. *The Journal of Agricultural Science*, *93*(3), 513–520. https://doi.org/10.1017/S0021859600038910

Suter, M., Connolly, J., Finn, J. A., Loges, R., Kirwan, L., Sebastià, M.-T., & Lüscher, A. (2015). Nitrogen yield advantage from grass–legume mixtures is robust over a wide range of legume proportions and environmental conditions. *Global Change Biology*, *21*(6), 2424–2438. https://doi.org/10.1111/gcb.12880

Thorburn, P. J., Biggs, J. S., Collins, K., & Probert, M. E. (2010). Using the APSIM model to estimate nitrous oxide emissions from diverse Australian sugarcane production systems. *Agriculture, Ecosystems and Environment*, *136*(3–4), 343–350. https://doi.org/10.1016/j.agee.2009.12.014

van Groenigen, J. W., Velthof, G. L., Oenema, O., van Groenigen, K. J., & Kessel, C. V. (2010). Towards an agronomic assessment of N_2O emissions: A case study for arable crops. *European Journal of Soil Science*, *61*(6), 903–913. https://doi.org/10.1111/j.1365-2389.2009.01217.x

Vellinga, T. V., de Haan, M. H. A., Schils, R. L. M., Evers, A., & van den Pol–van Dasselaar, A. (2011). Implementation of GHG mitigation on intensive dairy farms: Farmers' preferences and variation in cost effectiveness. *Livestock Science*, *137*(1), 185–195. https://doi.org/10.1016/j.livsci.2010.11.005

Virkajärvi, P., Maljanen, M., Saarijärvi, K., Haapala, J., & Martikainen, P. J. (2010). N_2O emissions from boreal grass and grass-clover pasture soils. *Agriculture, Ecosystems & Environment*, *137*(1–2), 59–67. https://doi.org/10.1016/j.agee.2009.12.015

Vogeler, I., & Cichota, R. (2016). Deriving seasonally optimal nitrogen fertilization rates for a ryegrass pasture based on agricultural production systems simulator modelling with a refined AgPasture model. *Grass and Forage Science*, *71*(3), 353–365. https://doi.org/10.1111/gfs.12181

Vogeler, I., Snow, V., Jolly, B., & Bryant, R. (2010). *Determining risk indicators for N leaching using APSIM modelling (AgResearch Internal Report)*. Palmerston North: AgResearch.

Well, R., & Butterbach-Bahl, K. (2010). Indirect emissions of nitrous oxide from nitrogen deposition and leaching of agricultural nitrogen. In K. A. Smith (Ed.), *Nitrous oxide and climate change* (pp. 162–189). London, UK:Earthscan.

White, R., Murray, S., & Rohweder, M. (2000). *Pilot analysis of global ecosystems: Grassland ecosystems*. Washington, DC, USA: World Resources Institute. Retrieved from https://www.wri.org/publication/pilot-analysis-global-ecosystems-grassland-ecosystems

Whitehead, D. C. (2000). *Nutrient elements in grassland: Soil-plant-animal relationships*. Wallingford, Oxon, UK: CABI Publishing.

Wilkins, P. W., Allen, D. K., & Mytton, L. R. (2001). Differences in the nitrogen use efficiency of perennial ryegrass varieties under simulated rotational grazing and their effects on nitrogen recovery and herbage nitrogen content. *Grass and Forage Science*, *55*(1), 69–76. https://doi.org/10.1046/j.1365-2494.2000.00199.x

Wood, S. N. (2006). *Generalized additive models: An introduction with R* (1st ed.) Boca Raton, FL: Chapman and Hall/CRC.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, *73*. https://doi.org/10.1111/j.1467-9868.2010.00749.x