Assessing the effects of agricultural management practices on crop ecosystems with the LPJ-GUESS model

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN (Dr. rer. nat.)
von der KIT Fakultät für
Bauingenieur-, Geo- und Umweltwissenschaften

des Karlsruher Instituts für Technologie (KIT)
genehmigte
DISSERTATION
von
Master of Meteorology: Jianyong Ma
aus Xinjiang, China

Tag der mündlichen Prüfung: 12.05.2023

Referentin: Prof. Dr. Almut Arneth
Korreferent: Prof. Dr. Sönke Zaehle

Karlsruhe 2023

Abstract

Substantial losses of global soil organic carbon (SOC) have occurred over recent decades, arising from agricultural intensification and the conversion of natural soils for agricultural uses to feed the growing population. Enhancing SOC stocks in croplands through improved management practices—such as reducing tillage, residue application and cover crops—has been identified as a promising option for climate change mitigation, with co-benefits for soil fertility and crop yields. However, the large-scale quantification of these management effects on agricultural ecosystems with ecological models remains uncertain, including the impacts of the cultivation of leguminous crops, which has the potential to reduce the use of synthetic fertilizer and therefore contributes to environmental sustainability. To better represent global agricultural managements, in this thesis I first incorporate two grain legumes (soybean and faba bean) and one herbaceous legume (white clover), with their inherent process of biological nitrogen fixation (BNF) into the dynamic global vegetation model LPJ-GUESS. The spatial and temporal patterns of the BNF rates in soybean and faba bean are quantified over the historical period 1981-2016. Subsequently, the large-scale influence of alternative management strategies (such as legume cover crops, no-tillage, and residue retention) on yields and cropland carbon (C) and nitrogen (N) balances is investigated under present and future climate conditions, through applying and analyzing results from the updated model version.

Simulations from LPJ-GUESS show that global N fixation in soybean and all pulses (representing faba bean in the model) are 11.6 ± 2.2 Tg N yr^{-1} and 5.6 ± 1.0 Tg N yr^{-1}, respectively, during the period 1981-2016. Spatially, the highest BNF rates are found in tropical and temperate regions with warm and humid climates. Soil water availability and temperature are most important factors controlling N fixation, in addition to N fertilization. Overall, the modelled total N fixation from grain legumes accounts for 12% of the annual N fixed across all global terrestrial ecosystems (ca. 140 Tg N yr^{-1}), indicating the importance of BNF input in croplands for the global terrestrial N cycle, although a large amount of the fixed N is removed from the ecosystems every year through crop harvests.

Practicing legume cover cropping would be quite different to the cultivation of grain legumes since the fixed N in cover crops is usually returned to soils. Assuming that all croplands worldwide are to adopt conservation agriculture techniques, the model estimates that combined N-fixing cover crops with no-tillage management can potentially increase soil carbon levels by

7% (+0.32 Pg C yr^{-1} in global croplands) while reducing N leaching loss by 41% (-7.3 Tg N yr^{-1}) after 36 years of implementation (the maximum duration found in cover cropping field trials in this thesis). This integrated practice is accompanied by a potential 2% increase in total crop production (+37 million tonnes per year including wheat, maize, rice, and soybean) in the last decade of the simulation. In comparison with non-legume cover crops, the adoption of N-fixing cover cropping in model experiments contributes more to gaining yield benefits in the humid tropics while it is mitigating production losses under northern temperate climates. This spatial variation is found to be associated with main-crop types and N fertilizer inputs, with little yield changes simulated in soybean systems and highly fertilized agricultural soils.

Taking eastern Africa as a case study, legume cover crops, together with six alternative management strategies, are further assessed to quantify their impacts on crop ecosystems. The regional simulations reveal that improved managements implemented in the warm and humid ecosystems can favor climate change mitigation while benefiting crop yields, especially for an integrated conservation agriculture practice that combines all soil-conserving techniques. When integrated the simulated grid cells over the study region, this combined strategy—including no-tillage, residue and manure application, and cover cropping—is found to increase total SOC stocks by 11% in the long term, accompanied by a 25% enhancement in total crop production under the present-day climate. Over the same period, practicing N-fixing cover crops is also identified as a promising strategy to increase cropland soil C levels (+4%) and agricultural production (+16%), however on the environmental cost of increased total N losses (+28%; including gaseous emissions and N leaching). These management influences would be possibly sustained in the future under three climate pathways.

In conclusion, the results of this thesis highlight the importance of accounting for N fixers when assessing large-scale C-N cycles in conservation agriculture systems. They also demonstrate that, with improved agricultural managements, it is possible to achieve environmental sustainability and ensure food security in global croplands.

Zusammenfassung

In den letzten Jahrzehnten ist es weltweit zu erheblichen Verlusten an organischem Kohlenstoff (SOC) im Boden gekommen, die auf die Intensivierung der Landwirtschaft und die Umwandlung natürlicher Böden in landwirtschaftliche Nutzflächen zur Ernährung der wachsenden Bevölkerung zurückzuführen sind. Die Erhöhung der SOC-Bestände in Ackerflächen durch verbesserte Bewirtschaftungspraktiken - wie die Verringerung der Bodenbearbeitung, die Ausbringung von Ernterückständen und der Anbau von Zwischenfrüchten - wurde als vielversprechende Option für die Eindämmung des Klimawandels identifiziert, mit gleichzeitigen Vorteilen für die Bodenfruchtbarkeit und die Ernteerträge. Die großflächige Quantifizierung dieser Bewirtschaftungspraktiken auf landwirtschaftliche Ökosysteme, einschließlich der Auswirkungen des Anbaus von Leguminosen, ist jedoch nach wie vor unsicher. Um die globale landwirtschaftliche Produktion besser abzubilden, integriere ich in dieser Arbeit zunächst zwei Körnerleguminosen (Sojabohne und Ackerbohne) und eine krautige Leguminose (Weißklee) mit biologischer Stickstofffixierung (BNF) in das dynamische Vegetationsmodell LPJ-GUESS. Die räumlichen und zeitlichen Muster der BNF-Raten in Sojabohnen und Ackerbohnen werden über den historischen Zeitraum 1981-2016 quantifiziert. Anschließend wird der Großflächige Einfluss alternativer Bewirtschaftungsstrategien auf die Ernteerträge und die Kohlenstoff- (C) und Stickstoff- (N) Bilanzen der Anbauflächen unter gegenwärtigen und zukünftigen Klimabedingungen untersucht, indem die Ergebnisse der aktualisierten Modellversion angewendet und analysiert werden.

Die Modellsimulationen zeigen, dass die globale N-Fixierung in Sojabohnen und allen Hülsenfrüchten (die im Modell die Ackerbohne repräsentieren) im Zeitraum 1981-2016 bei 11,6 $\pm 2,2$ Tg N yr^{-1} bzw. 5,6$\pm$1,0 Tg N yr^{-1} beträgt. Räumlich gesehen sind die höchsten BNF-Raten in tropischen und gemäßigten Regionen mit warmem und feuchtem Klima zu finden. Die Bodenwasserverfügbarkeit und die Temperatur sind neben der N-Düngung die wichtigsten Einflussfaktoren für die N-Fixierung. Insgesamt macht die modellierte Gesamt-N-Fixierung durch Körnerleguminosen 12 % des jährlich in globalen terrestrischen Ökosystemen fixierten N aus (ca. 140 Tg N yr^{-1}), was auf die Bedeutung des BNF-Eintrags in Ackerflächen für den globalen terrestrischen N-Kreislauf schließen lässt, obwohl ein großer Teil des fixierten N jedes Jahr durch die Ernte aus den Ökosystemen entfernt wird.

Der Anbau von Leguminosen als Deckfrucht in der Zwischenseason unterscheidet sich deutlich vom reinen Anbau von Körnerleguminosen, da der in Deckfrüchten fixierte Stickstoff in der Regel in den Boden zurückgeführt wird. Unter der Annahme, dass weltweit alle Anbauflächen konservierende Landwirtschaftstechniken verwenden, ergibt sich basierend auf den Modelldaten, dass die Kombination von N-fixierenden Deckfrüchten und minimaler Bodenbearbeitung den Kohlenstoffgehalt des Bodens um 7 % (+0,32 Pg C yr^{-1} in den globalen Anbauflächen) erhöhen und gleichzeitig die N-Auswaschungsverluste um 41 % (-7,3 Tg N yr^{-1}) nach 36 Jahren der Umsetzung reduzieren kann (die maximale Dauer, die in Feldversuchen mit Deckfrüchten in dieser Dissertation ermittelt wurde). Diese integrierte Praxis geht mit einem Anstieg der gesamten pflanzlichen Produktion um 2 % (+37 Millionen Tonnen pro Jahr, einschließlich Weizen, Mais, Reis und Soja) im letzten Jahrzehnt der Simulation einher. Im Vergleich zu Nicht-Leguminosen-Deckungskulturen trägt der Einsatz von N-fixierendem Deckungsanbau in den Modellexperimenten stärker zur Ertragssteigerung in den feuchten Tropen bei, während die Produktionsverluste in den nördlichen gemäßigten Klimazonen gemildert werden. Diese räumliche Variation hängt mit den Hauptkulturen und dem Stickstoffdüngereinsatz zusammen, wobei bei Sojabohnensystemen und stark gedüngten landwirtschaftlichen Böden nur geringe Ertragsveränderungen simuliert werden.

Am Beispiel von Ostafrika werden Leguminosen zusammen mit sechs alternativen Bewirtschaftungsstrategien untersucht, um ihre Auswirkungen auf die Ökosysteme von Nutzpflanzen zu quantifizieren. Die regionalen Simulationen zeigen, dass die verbesserten Bewirtschaftungsmethoden, die in den tropischen Ökosystemen umgesetzt werden, den Klimawandel abmildern und gleichzeitig die Ernteerträge steigern können, insbesondere bei einer integrierten konservierenden Landwirtschaft, die alle bodenschonenden Techniken kombiniert. In den untersuchten Regionen führt diese kombinierte Strategie, die keine Bodenbearbeitung, die Ausbringung von Rückständen und Dung sowie den Anbau von Deckfrüchten umfasst, langfristig zu einer Erhöhung der simulierten SOC-Vorräte um 11 %, begleitet von einer Steigerung der gesamten Pflanzenproduktion um 25 %. Der Anbau von N-fixierenden Deckfrüchten ist ebenfalls vielversprechend, um den C-Gehalt im Ackerboden (+4 %) und die landwirtschaftliche Produktion (+16 %) zu erhöhen, wobei die Umweltkosten in Bezug auf die gesamten N-Verluste (+28 %; einschließlich gasförmiger Emissionen und N-Auswaschung) zu berücksichtigen sind. Diese Bewirtschaftungseinflüsse würden bei drei Klimaszenarien möglicherweise auch in Zukunft bestehen bleiben.

Zusammenfassend zeigen die Ergebnisse dieser Arbeit, wie wichtig die Berücksichtigung von N-Fixierern bei der Bewertung großräumiger C-N-Zyklen in Systemen der konservierenden Landwirtschaft ist. Sie zeigen auch die Möglichkeit einer verbesserten landwirtschaftlichen Bewirtschaftung auf, um ökologische Nachhaltigkeit zu erreichen und die Ernährungssicherheit in globalen Anbauflächen zu gewährleisten.

Eidesstattliche Versicherung gemäß § 13 Absatz 2 Satz 1 Ziffer 4 der Promotionsordnung des Karlsruher Instituts für Technologie (KIT) für die KIT-Fakultät für Bauingenieur-, Geo- und Umweltwissenschaften.

1. Bei der eingereichten Dissertation zu dem Thema "Assessing the effects of agricultural managment practices on crop ecosystems with the LPJ-GUESS model" handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts ver-schwiegen habe.

Ort und Datum   Unterschrift

This thesis is submitted as a cumulative dissertation and consists of three main chapters (Chapters 3-5). All of these have been published in peer-reviewed journals. The chapters are as follows:

3. Assessment of biological nitrogen fixation in global grain legumes

This chapter is based on the paper: <u>Ma, J.</u>, Olin, S., Anthoni, P., Rabin, S. S., Bayer, A. D., Nyawira, S. S., & Arneth, A. (2022). Modeling symbiotic biological nitrogen fixation in grain legumes globally with LPJ-GUESS (v4.0, r10285). *Geoscientific Model Development*, *15*(2), 815–839. https://doi.org/10.5194/gmd-15-815-2022.

4. Global influence of cover crops on yield and cropland carbon and nitrogen balances

This chapter is based on the paper: <u>Ma, J.</u>, Anthoni, P., Olin, S., Rabin, S. S., Bayer, A. D., Xia, L., & Arneth, A. (2023). Estimating the global influence of cover crops on ecosystem service indicators in croplands with the LPJ-GUESS model. *Earth's Future*, *11*(5), e2022EF003142. https://doi.org/10.1029/2022EF003142.

5. Impacts of agricultural management practices on soil carbon stocks, nitrogen loss, and crop production in eastern Africa

This chapter is based on the paper: <u>Ma, J.</u>, Rabin, S. S., Anthoni, P., Bayer, A. D., Nyawira, S. S., Olin, S., Xia, L., & Arneth, A. (2022). Assessing the impacts of agricultural managements on soil carbon stocks, nitrogen loss and crop production — a modelling study in eastern Africa. *Biogeosciences*, *19*(8), 2145–2169. https://doi.org/10.5194/bg-19-2145-2022.

Due to the papers being published, and therefore involving the work of co-authors, I detail my contribution to the Chapters 3-5 as follows:

3. I implemented biological nitrogen fixation process in grain legumes into the LPJ-GUESS model, based on an original suggestion by Dr. Stefan Olin, designed the experiments, collected the observational data for evaluation, and led the analysis and the writing of the paper.

4. I implemented biological nitrogen fixation process in a herbaceous legume into the LPJ-GUESS model, based on an original suggestion by Dr. Stefan Olin, designed the experiments, collected the observational data for evaluation, and led the analysis and the writing of the paper.

5. I designed the experiments, processed the model input data, performed the simulations and analysis, and led the writing of the paper.

Contents

Abbreviations and units

%Ndfa	The proportion of plant nitrogen derived from the atmosphere
Ag-GRID	The AgMIP Gridded Crop Modeling Initiative
APSIM	The Agricultural Production Systems sIMulator model
BNF	Biological Nitrogen Fixation
C	Carbon
C:N	Carbon-to-nitrogen ratio
CA	Conservation Agriculture
CA_{all}	All global croplands under conservation agriculture practice
CA_{his}	Current conservation agriculture areas in global croplands
CA_{pot}	Potential croplands that may adopt conservation agriculture practice
CC	Cover Crop
CC_L	Management scenario of legume cover crop
CC_LNT	Management scenario of combining legume cover crop with no-tillage
CCMI	Chemistry-Climate Model Initiative
CC_{NL}	Management scenario of non-legume cover crop
CFT	Crop Functional Type
CH_4	Methane
CO_2	Carbon dioxide
CO_2eq	Carbon dioxide equivalent
CONSERV	Management scenario of the integrated conservation agriculture practice
CROPGRO	Crop growth simulator
CRUJRA	Climatic Research Unit and Japanese reanalysis data
DGVM	Dynamic Global Vegetation Model
DM	Dry Matter
DS	Crop development stage
EPIC	Erosion/Productivity Impact Calculator model
FAO	Food and Agriculture Organization of the United Nations
GCM	General Circulation Model
GFDL-ESM4	Geophysical Fluid Dynamics Laboratory Earth System Model version 4
Gg	Giga-gram ($1Gg = 10^9$ grams)

GGCMI	Global Gridded Crop Model Intercomparison
GHG	Greenhouse gas
GPP	Gross Primary Production
GSWP3-W5E5	Global Soil Wetness Project Phase 3-WFDE5 over land merged with ERA5 over the ocean
GtC	Giga-ton of carbon (1 Gt=10^{15} grams)
IPSL-CM6A-LR	Climate model of the Institut Pierre Simon Laplace-low resolution
ISIMIP	Inter-Sectoral Impact Model Intercomparison Project
ITPS	Intergovernmental Technical Panel on Soils
LPJ-GUESS	Lund-Potsdam-Jena General Ecosystem Simulator
LPJmL	Lund-Potsdam-Jena model with managed Land
LUC	Land-Use Change
LUH2	Land-Use Harmonization v2
MAE	Mean Absolute Error (%)
MAN	Management scenario of manure application
ME	Mean Error (%)
Mg	Mega-gram (1 Mg = 10^6 grams)
Mha	Million hectares (1 Mha = 10^6 ha)
MIRCA	Global data set of monthly irrigated and rainfed crop areas around the year 2000
MPI-ESM1-2-HR	Max Planck Institute Earth System Model version 1-high resolution
MRI-ESM2-0	Meteorological Research Institute Earth System Model version 2
N	Nitrogen
N_2O	Nitrous oxide
NH_3	Ammonia
NoCC	Management scenario of no cover crop
NPP	Net Primary Production
NT	Management scenario of no-tillage
OPT	Management scenario of the optimal practice for soil carbon stocks
PFT	Plant Functional Type
Pg	Peta-gram (1 Pg = 10^{15} grams)
PHU	Potential Heat Unit (degree-day)

PNV	Potential Natural Vegetation
ppm	Parts per million
r	Pearson correlation coefficient
RCP	Representative Concentration Pathway
RMSE	Root Mean Square Error
RR	Management scenario of residue retention
SLA	Specific Leaf Area
SOC	Soil Organic Carbon
SOILN	Soil nitrogen simulator
SOM	Soil Organic Matter
SSA	Sub-Saharan Africa
SSP	Shared Socioeconomic Pathway
STD	Management scenario of standard simulation
STICS	Multidisciplinary simulator for standard crops
Tg	Tera-gram (1 Tg = 10^{12} grams)
UKESM1-0-LL	United Kingdom Earth System Model version 1-low resolution
ΔSOC_{rate}	Annual soil organic carbon change rate (mega-gram of carbon per year)

List of Figures

List of Tables

1 General introduction

Agricultural intensification and the continuous conversion of natural soils (e.g., forests and grasslands) for agricultural uses to feed the growing human population has until now been the dominant drivers of soil degradation and other significant side effects on the environment. Improved agricultural management practices, such as reducing tillage, manure application, or leguminous cover crops, are expected to enhance soil fertility, with climate change mitigation co-benefits while increasing crop production. However, the large-scale quantification of crop management effects on agricultural ecosystems remains uncertain. In this thesis, I investigate the influence of agricultural practices on yields and cropland carbon and nitrogen processes by applying and analyzing results from a process-based dynamic global vegetation model. The introduction of the thesis starts with an overview about perturbations of the terrestrial carbon cycle and climate through land-use changes, followed by a closer examination of agricultural land use and its potential impacts on environmental sustainability. Afterwards, I provide some background information about agricultural practices aimed at managing soil fertility for improved food security and climate change mitigation. Lastly, I introduce recent progress in global vegetation modelling in terms of the improved agricultural managements before describing the structure and research questions addressed in this thesis.

1.1 Impacts of land-use changes on the terrestrial carbon cycle and climate

Most of the carbon (C) on our planet is stored in the lithosphere as sedimentary carbonates (66 000 000 – 100 000 000 GtC; 1 Gt = 10^{15} grams), followed by the oceans, mainly in the form of dissolved inorganic carbon at great depths (37 000 GtC). A much smaller proportion is stored in rock formations as fossil fuel deposits (4000 GtC), or exists in soil organic matter (SOM; 1700 GtC), the atmosphere (875 GtC), and terrestrial plants (450 GtC; see Fig. 1.1). Although these carbon pools seem negligible compared with the stocks in the lithosphere and ocean, carbon exchanges among Earth's spheres, mainly in the form of carbon dioxide (CO_2), take place much faster in the small pools, and the larger reservoirs are less involved with the C cycle on a human timescale. Each year (2010-2019 average), the terrestrial vegetation is estimated to take up 113 GtC from the atmosphere (gross primary productivity, GPP; Canadell et al., 2021) via the photosynthetic reduction of CO_2 into organic compounds (i.e., natural process without anthropogenic perturbation). The "fixed" C from the atmosphere is partially used for plant autotrophic respiration, with the remainder (net primary productivity, NPP) allocated to structural biomass in stems, leaves, roots,

and fruits for supporting plant growth. After the death of the plant, the C in residues is transferred to the aboveground and belowground litter pools and released back into the atmosphere as CO_2 through the decomposition of soil microorganisms (heterotrophic respiration). Currently, global total GPP is estimated to exceed global respiration and C emissions from natural disturbances (e.g., wildfire), the terrestrial biosphere thus acts as a net C sink, with a mean uptake rate of 3.1 ± 0.6 GtC yr^{-1} (2012-2021 average; Friedlingstein et al., 2022).
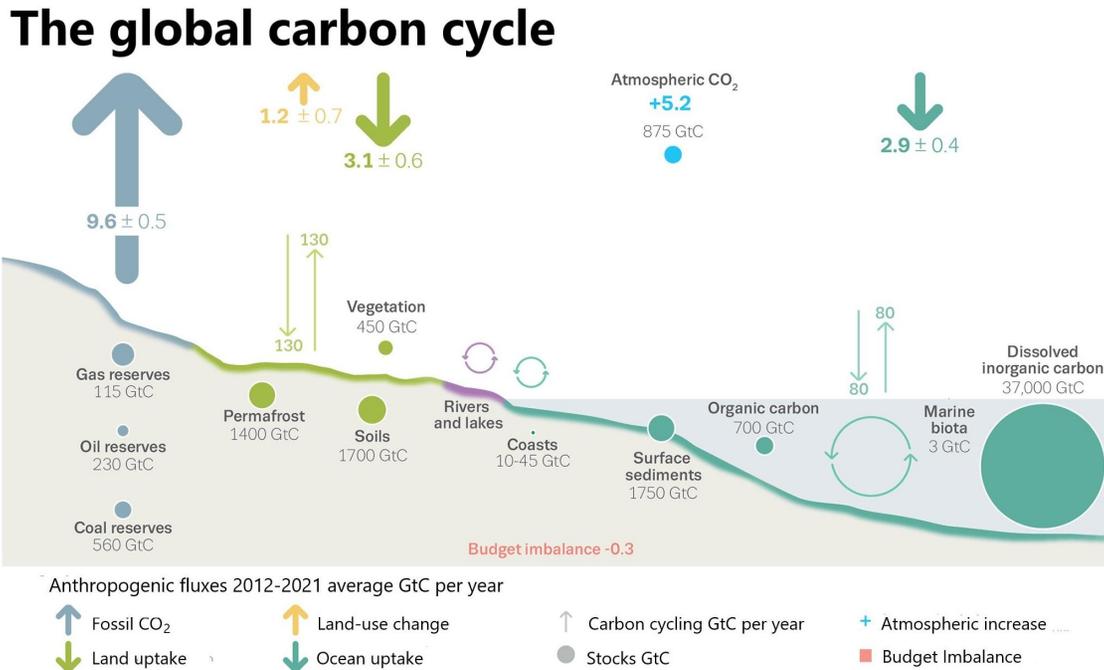
Figure 1.1 Schematic representation of the global carbon cycle averaged over 2012-2021. The solid circles and arrows represent reservoir stocks and exchange fluxes, respectively, with values given in the background. The figure is taken from Friedlingstein et al. (2022).

Since the dawn of the Industrial Revolution, human activities have dramatically altered the C cycle, mainly by burning fossil fuels (e.g., coal, oil, and gas) and cement production, as well as by land-use change (LUC). From 1850 to 2021, cumulative CO_2 emissions from fossil fuel combustion and LUC are estimated to be 465 ± 25 GtC and 205 ± 60 GtC, respectively, with an average of 9.6 ± 0.5 GtC yr^{-1} and 1.2 ± 0.7 GtC yr^{-1} for the decade 2012-2021 (see Fig. 1.2). As a result, atmospheric CO_2 concentration has significantly increased by over 40% from a pre-industrial level of 280 ppm to 416 ppm in 2022 (Lan et al., 2023). The conversion of natural forests for agricultural use is the dominant driver of global LUC emissions (Friedlingstein et al., 2022) through releasing the biomass C stored in forests into the atmosphere either directly (forests are burned) or subsequently (wood products are burned or decomposed). In recent decades, significant deforestation happened for

soybean production in Brazil (Heilmayr et al., 2020) and land clearing for large-scale oil palm plantation in Indonesia (Austin et al., 2019). On average over the period of 2012-2021, the estimates of global C loss from deforestation amount to 1.8 ± 0.4 GtC yr^{-1}, two times higher than the C gain of 0.9 ± 0.3 GtC yr^{-1} due to re/afforestation (Friedlingstein et al., 2022). Converting forests into agricultural lands usually leads to continued soil organic carbon (SOC) losses since crop harvest removes carbon out of the ecosystem and usually only a small portion is returned to the soil. In addition, erosion/leaching rates and heterotrophic respiration on croplands are usually higher than in natural vegetation in general (Smith et al., 2016). Estimates suggest that converting forests to croplands has reduced SOC stocks by 32-36% in temperate (Poeplau et al., 2011) and 25-30% in tropical regions (Don et al., 2011), with cumulative SOC losses ranging from 133-186 GtC across the globe since 1850 (Lal, 2004; Smith et al., 2016; Sanderman et al., 2017).

LUC can influence local to global-scale climate via altering atmospheric composition (i.e., biogeochemical changes) and the flow of energy and water between the land and the atmosphere (i.e., biophysical changes). Increases in atmospheric CO_2 concentration belong to biogeochemical changes, with other effects being emissions of methane (CH_4; mainly from ruminants and rice production) and nitrous oxide (N_2O; mainly from the intensive use of synthetic fertilizer). In contrast to global biogeochemical climate change impacts, biophysical changes (e.g., albedo, evapotranspiration, and surface roughness) arising from LUC mainly impact the climate on the local to regional scales (Arora & Montenegro, 2011; Peng et al., 2014; Li et al., 2020). For example, several studies suggest that forests tend to be cooler than herbaceous croplands throughout much of the tropics, and reforestation in the tropical forests would promote cooling regionally (Anderson et al., 2011; Mildrexler et al., 2011; Li et al., 2015). These cooling effects in general decrease with increasing latitude, but the magnitude of the cooling from reforestation is highly dependent on the specific locations and the season of the year (Wickham et al., 2014; Zhao & Jackson, 2014; Li et al., 2015). Irrigated agriculture is another big factor affecting global carbon, water and nutrient cycles. Currently, irrigated agriculture accounts for 22% of cultivated lands (FAOSTAT, 2023), consuming 85-90% of global anthropogenic freshwater use (D'Odorico et al., 2018; Rosa, 2022). Such a substantial water application to croplands can affect climate as well. A modelling study by Gordon et al. (2005) estimated that cropland irrigation enhanced global water vapor flows to the atmosphere by 2600 km^3 yr^{-1}—an increase that was approximately as large as the decreased vapor flows caused by deforestation (3000 km^3 yr^{-1}). Additionally, due to irrigation-induced increases both in cloud

cover (Lobell et al., 2009; Sacks et al., 2009) and downwind precipitation (DeAngelis et al., 2010), irrigated agriculture was found to significantly reduce near-surface air temperatures in some regions (e.g., northern mid-latitudes and portions of south Asia), although it showed a negligible cooling effect in terms of global average (Sacks et al., 2009; Sleeter et al., 2018).

Anthropogenic carbon flows
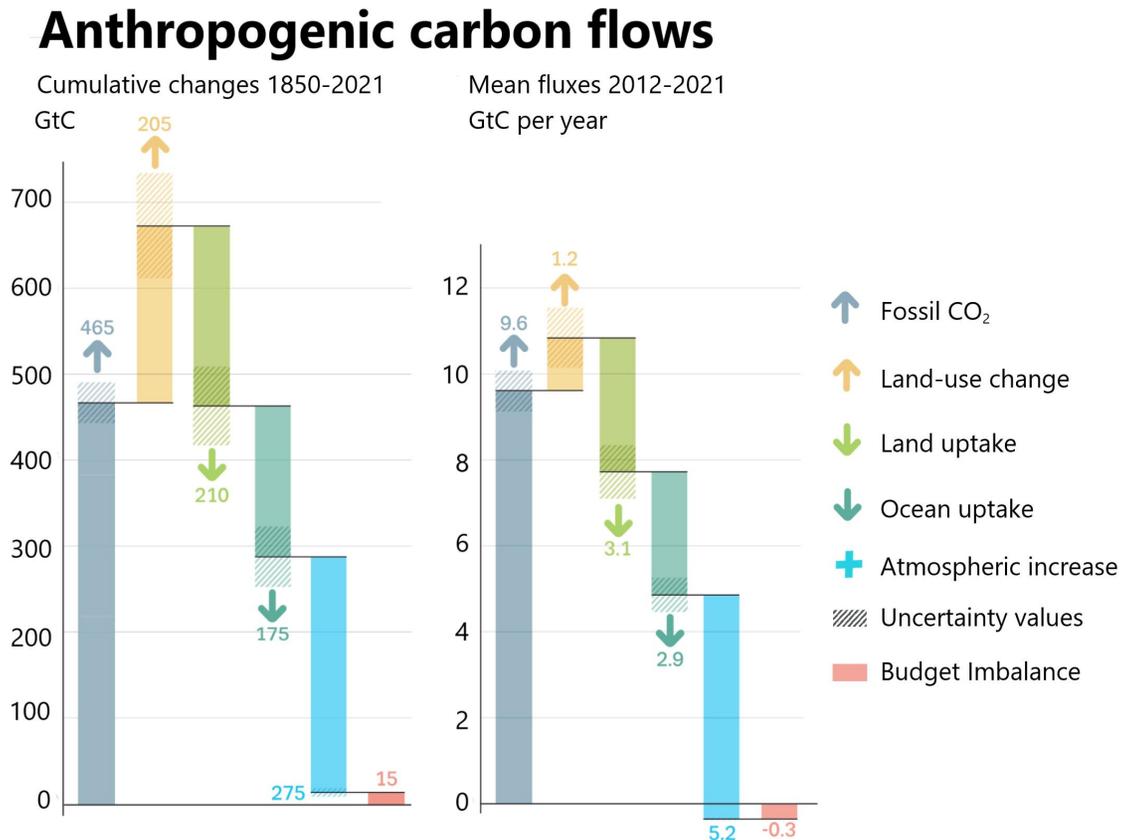
Figure 1.2 Cumulative carbon emissions during the 1850-2021 period (left) and mean annual fluxes over 2012-2021 (right) caused by the anthropogenic perturbation of the global carbon cycle. The figure is taken from Friedlingstein et al. (2022).

1.2 Agricultural land use and its impacts on environmental sustainability

In year 2020 global agricultural land area was roughly 4700 million hectares (Mha), accounting for 37% of the Earth's land surface (FAOSTAT, 2023). About one-third of this is used for the production of crops (i.e., cropland), while the remaining two-thirds are mostly used as pasture for grazing livestock. As the global population continues to grow, there is increasing demand for food and therefore land. According to the estimates by Hurtt et al. (2020), the global land surface used for cropland has increased to 1560 Mha in 2014, compared with pre-industrial agricultural area with 560 Mha in 1850 (Land-Use Harmonization datasets v2—LUH2; see Fig 1.3). In particular, the harvested area for soybean more than quadrupled from 24 Mha in 1961 to 127 Mha in 2020

(FAOSTAT, 2023; see Fig. 1.3) as a consequence of the high protein content and attractive cash return from its grain yield. There are substantial concerns about the sustainability of soybean production in the last two decades, due to its links to deforestation and loss of native vegetation especially in South America (Heilmayr et al., 2020; Song et al., 2021). According to the estimates by Hurtt et al. (2020), the present-day deforestation for cropland expansion is likely to continue over the next decades in most future scenarios (see Fig. 1.3), as a result of the need to feed 8.4-9.9 billion people by 2050 (Samir & Lutz, 2017; Beltran-Peña et al., 2020). However, in addition to causing C losses in vegetation and soils (West et al., 2014; Friedlingstein et al., 2022), increasing food production through agricultural extensification also has other detrimental effects, such as the decline of the Earth's biodiversity (Green et al., 2019; Arneth et al., 2020). Estimates suggest that LUC has caused ecological assemblages to lose on average 13.6% of global terrestrial species in comparison with pristine habitats, with the worst-affected habitats losing 76.3% (Newbold et al., 2015; Newbold, 2018).

As an alternative to agricultural extensification, crop production can be increased by increasing yield on existing croplands via improved technology and management (also known as agricultural intensification; Rosa, 2022). Since water and nutrients are major limiting factors constraining crop growth (Mueller et al., 2012), irrigation and fertilizer application are two effective strategies to enhance crop yields worldwide (Beltran-Peña et al., 2020). Considering the fact that irrigated agriculture often already depletes freshwater resources at massive scales (D′Odorico et al., 2018) and that the scarcity of irrigation water can rarely be solved by physical water transfers over long distances, it is difficult to close yield gaps—defined as the difference between the theoretical maximum possible yield level and actual farmers' yield (Van Ittersum et al., 2013)—in dry regions through expanding irrigation on currently rain-fed croplands. In contrast, nutrient limitation to crop productivity can be overcome through fertilizer application wherever nutrients are constraining yields (Erisman et al., 2008). As a result, anthropogenic N fertilizer inputs on agricultural lands have increased nearly six-fold between the 1960s and 2010s, reaching 96.5 Tg N yr^{-1} (1 Tg = 10^{12} grams) and 8.5 Tg N yr^{-1} for cropland and pasture, respectively, in the 2010s (Tian et al., 2022). Nevertheless, estimates indicate that crops' global average nitrogen use efficiency—the fraction of N input harvested as product—is only around 42% (Zhang et al., 2015; Xia & Yan, 2023) and that about half of N investments to agricultural soils are lost to the environment (Sutton et al., 2011; Zhang et al., 2015; Gu et al., 2023) through gaseous emissions (e.g., N_2O and NH_3) and

hydrological processes (e.g., leaching and runoff), contributing to climate change, severe air and water pollution, and soil acidification.
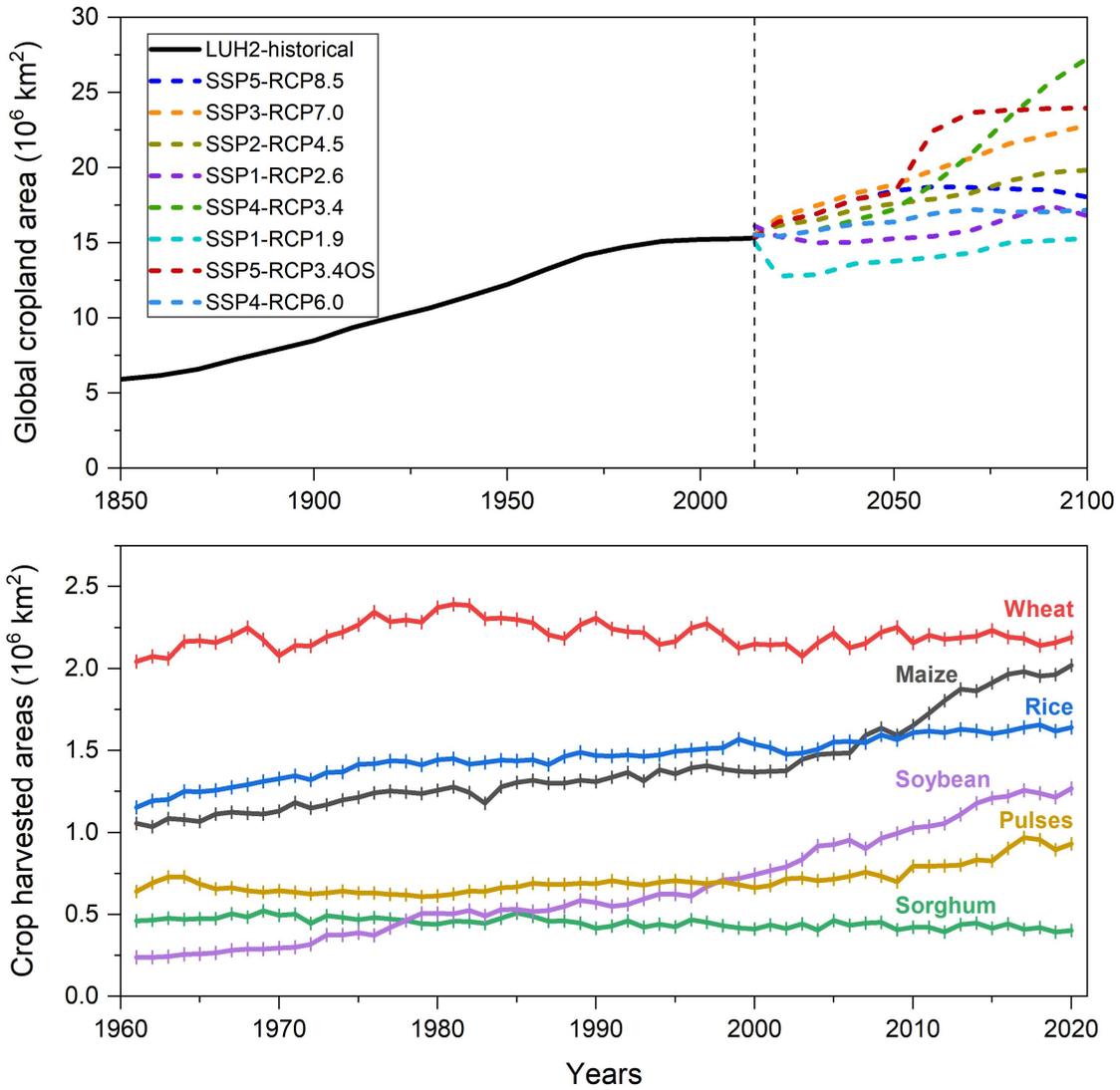
Figure 1.3 Estimates of global cropland areas over 1850-2014 and 2015-2100 under eight future scenarios (top; Hurtt et al., 2020), and FAO-reported global harvested areas of six crops for the 1961-2020 period (down; FAOSTAT, 2023). The top figure is replotted based on the data from Hurtt et al. (2020).

At present, global human-induced N_2O emission is estimated at 7.3 Tg N yr^{-1} (2007-2016 average), and about 52% of those anthropogenic emissions come from the agricultural sector (3.8 Tg N yr^{-1}; Tian et al., 2020). Livestock manure (including manure left on pastures by grazing animals and manure applied to croplands) and synthetic N fertilizer are the biggest drivers of agricultural N_2O emissions; these two activities also contribute 20% and 13% to annual total non-CO_2 greenhouse gas (GHG) emissions, respectively (FAO, 2020; see Fig. 1.4). The rapid increase of atmospheric

N_2O concentrations not only contributes to increasing near-surface air temperature (N_2O is about 300 times as potent as CO_2 at warming the atmosphere; Canadell et al., 2021), but also depletes stratospheric ozone, thus increasing surface levels of harmful ultraviolet radiation (Prather et al., 2015; Revell et al., 2015).
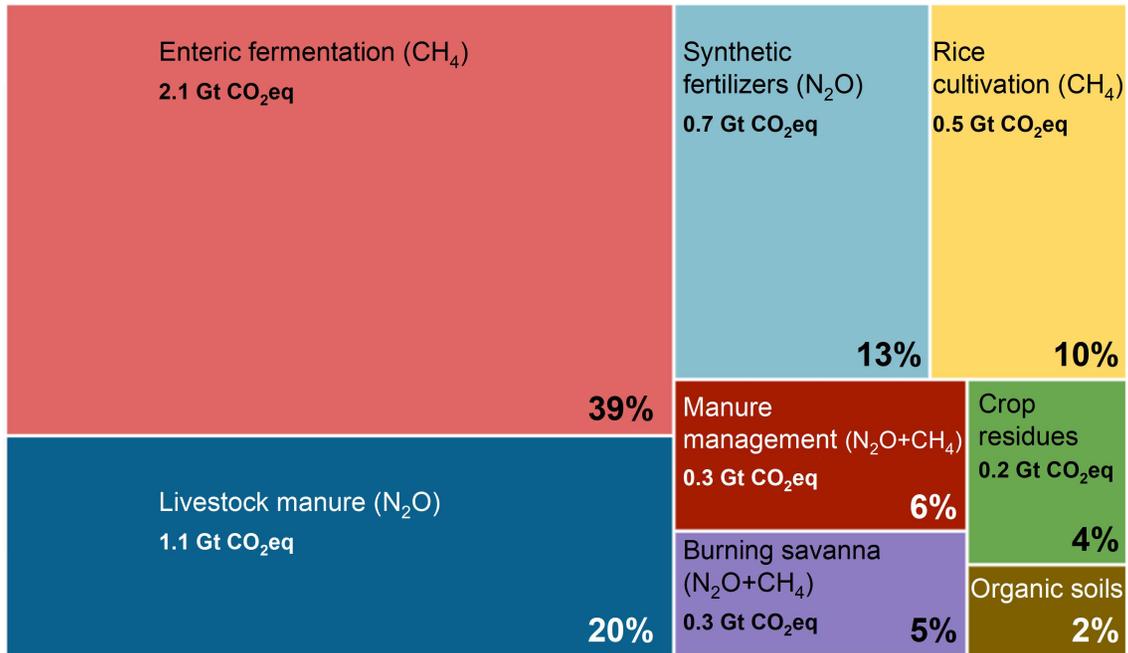
Figure 1.4 Contribution of crops and livestock activities to total non-CO_2 greenhouse gas (GHG) emissions from agriculture in 2018 (5.3 Gt CO_2eq). The figure is replotted based on the data reported in FAO (2020).

In addition to enhancing the abundance of N_2O in the atmosphere, livestock waste and the overuse of N fertilizers on agricultural soils are major contributors to global NH_3 volatilization (Liu et al., 2022), a key precursor of secondary aerosol (especially fine particle matter; $PM_{2.5}$) that have adverse impacts on air quality and human health (Shen et al., 2020; Vira et al., 2022). Over the past four decades, global agricultural NH_3 emissions has increased by more than 70% from 36 Tg N yr^{-1} in 1980 to 64 Tg N yr^{-1} in 2018, in which NH_3 emissions caused by synthetic fertilizers and livestock manure have increased by 128% (14-32 Tg N yr^{-1}) and 45% (22-32 Tg N yr^{-1}), respectively (Liu et al., 2022). The combination of excessive NH_3 deposition and nitrate leaching, arising from the surplus N in fertilized soils, is harmful to terrestrial and freshwater systems, and drive soil acidification (Steffen et al., 2015; Shen et al., 2020), eutrophication (Moss, 2008), and biodiversity loss (Erisman et al., 2013; Gu et al., 2023). To make matters worse, acidified soils and the continuous SOC loss on croplands (primarily from deforestation, see Sect. 1.1) are accompanied by low agricultural yields due to land degradation (Lal, 2004; Poeplau & Don, 2015). Accelerated

soil degradation has been reported on as much as 500 Mha in the tropical region (Lamb et al., 2005), and a full 90% of the land's topsoil (i.e., ice-free land surface) is likely to become degraded by 2050 (FAO & ITPS, 2015). Considering the huge food demand to sustain the population growth in the next decades (Samir & Lutz, 2017; Beltran-Peña et al., 2020) and the ongoing shift to richer diets (Tilman et al., 2011), it is crucial to improve agricultural productivity in order to ensure food security while reducing cropland C and N losses and restoring degraded soils for environmental sustainability (Smith et al., 2020; Arneth et al., 2021).

1.3 Alternative cropland management practices for improved soil carbon sequestration

Restoring degraded agricultural soils and enhancing soil carbon brings benefits to all farms and to society more generally. The benefits include increased soil water retention (resilience for rain-fed agriculture), improved soil nutrient potential to maintain long-term agricultural productivity, and increased C sequestration to reduce GHG emissions for climate change mitigation (Lal, 2015; Arneth et al., 2021). Given all of the benefits, interests in soil organic carbon, the major indicator of soil fertility, is now growing through various international initiatives (e.g., ARF100, 4per1000, ProSoil). For example, launched by the French government during COP 21 in 2015, the 4per1000 initiative sets a target of 3.4 Pg C yr^{-1} (1 Pg = 10^{15} grams) SOC sequestration in global agricultural topsoil (0-40cm) to mitigate climate change and support food security through the implementation of alternative management strategies (such as agroecology, agroforestry, conservation agriculture). As mentioned earlier, this thesis mostly explores the interactions between crop ecosystems and the atmosphere, I thus introduce the background information of the improved management practices on croplands usually adopted by farmers in the following.

The imbalance between carbon inputs (e.g., plant residue and manure application) and outputs (e.g., through crop harvest, decomposition of residues, leaching, and soil erosion) drives SOC storage changes in croplands. Conservation agriculture (CA)—in particular the adoption of minimum soil disturbance (e.g., no- or reduced tillage), permanent soil organic cover (e.g., crop residue returned to the soil, and cover crops), and species diversification through varied crop sequencing—is the most well-known and widely accepted practice to potentially increase soil carbon levels and enhance agricultural sustainability globally (Smith et al., 2016; Zomer et al., 2017). At present the world's CA area is estimated at 110 Mha, approximately amounting to 7-10% of global croplands (Porwollik et al., 2019). Much experimental evidence has indicated that SOC stocks under no-till systems are significantly higher than conventional farming practices due to the reduced

decomposition rate of soil organic matter (Pittelkow et al., 2015; Powlson et al., 2016; Sanderman et al., 2017; Sommer et al., 2018). However, the SOC benefits of no-till farming are statistically significant only in the topsoil (0-15 cm) and decline with soil depth (Haddaway et al., 2017). In global meta-analyses (Luo et al., 2010; Powlson et al., 2014), SOC stocks under no-till cropping systems were sometimes found to even be lower than conventional tillage in the deeper soil layers (>30 cm). Increasing organic material inputs with high C content as soil amendments are thus expected to be an alternative management practice for achieving SOC enhancement (Poeplau & Don, 2015). Retaining crop residue in the fields after harvest—either left on the soil surface or incorporated to the soils—has for many years been recommended as an important strategy in CA systems to strengthen resistance of SOC in croplands to soil degradation caused by intensive agriculture (Lal, 2004). The total crop residue production in the world was estimated as much as 3.8 Pg dry matter yr^{-1} for 27 food crops in the 2000s (Lal, 2005), and SOC content on croplands has the potential to increase by 9-13% if all harvested crop residue is returned to the soils (Xia et al., 2018; Bolinder et al., 2020). However, practical challenges still exist in terms of the wide implementation of crop residue retention on smallholder farms (Corbeels et al., 2014). In mixed crop-livestock systems, there is competition for residues between mulching and livestock for feed, particularly in African countries, where farmers also use residues for fuel and building (Thierfelder et al., 2013). In addition, short-term yield effects have been found to be variable: there are examples of no yield benefits and even short-term yield reductions after adoption of no-tillage and residue management (Corbeels et al., 2014; Stevenson et al., 2014; Pittelkow et al., 2015). Farmers would suffer economic losses, although it was estimated that these CA practices had a potential to increase global cropland SOC storage by 0.9-1.85 Pg C annually after more than 20 years of implementation (Zomer et al., 2017).

Application of livestock manure in agricultural land is also an important and widespread practice in view of enhancing SOC sequestration and improving crop yields due to the additional C and N inputs to the soils (Maillard & Angers, 2014). Over the past 150 years, global manure N production has increased by more than nine-fold from 9.5 Tg N yr^{-1} in the 1860s to 98.3 Tg N yr^{-1} in the 2010s, and is now similar to the amount of synthetic fertilizers applied to agricultural soils (105 Tg N yr^{-1}; Tian et al., 2022). Cattle (including dairy cows) dominated the manure N production among various livestock species, accounting for over 40% of total manure N production in the 2010s, followed by goats, sheep, and swine. Poultry (e.g., chickens and ducks) contributed the least to manure N compaction, they only played an important role in Canada and Russia (Zhang et al., 2017). Manure

might increase carbon stocks in soils as it has high carbon content or high C:N ratio. In global meta-analyses (Maillard & Angers, 2014; Gross & Glaser, 2021), the application of manure on average enhanced SOC sequestration on agricultural soils by 9.4-10.7 Mg C ha^{-1} (1 Mg = 10^6 grams) compared with the unfertilized treatments, or equivalent to a 35% increase in soil carbon stocks. Additionally, cropland SOC increases in response to manure use were found to be proportional to the cumulative manure-C input, but the scale of these positive effects varied widely across the region due to the differences in local climates, soil properties, manure application rates, as well as in manure types (considerable variation in C:N ratio arising from different plant species consumed by the farm animals; Maillard & Angers, 2014). Large-scale studies from regional to global levels are thus needed to comprehensively quantify the responses of SOC stocks to manure application.

Cover crops (CCs), also known as catch crops, are plants that mostly grow during the fallow period and are incorporated into the soils as "green manure" before planting the subsequent main crop. Experimental evidence has indicated that CC cultivation within agricultural rotations may increase SOC stocks by 13.8-17.3% over a period of up to 54 years, compared with management in which the off-season is left fallow (Poeplau & Don, 2015; Jian et al., 2020). Replacing bare fallows with cover cropping was estimated to sequester 0.16 Pg C yr^{-1} of soil carbon if 15% of global croplands were to adapt CC practice (Jian et al., 2020), which is 13% of current LUC emissions (1.2 Pg C yr^{-1}), or equivalent to about 2% of carbon emissions from fossil fuels (9.6 Pg C yr^{-1}; see Fig. 1.2). In addition to increasing organic matter inputs, CCs are also able to take up excess N from the soil and thus reduce N leaching (Tonitto et al., 2006; Thapa et al., 2018; Nouri et al., 2022; Blanchy et al., 2023), as well as to prevent the soils from compaction and erosion that happen when soils are bare (Kaye & Quemada, 2017). Moreover, compared with non-legumes, using legume CCs as green manure is more effective to maintain and/or improve soil fertility and crop productivity, since they not only increase soil organic matter via returning their biomass to soils, but also bring additional N into the soil as a result of their symbiotic association with rhizobial bacteria in root nodules—a process called biological nitrogen fixation (BNF). In general, the production of grain legumes increases linearly with legume BNF rate (Unkovich et al., 2010; Córdova et al., 2019), and the N benefit to soil fertility from green manure is strongly determined by N fixation capacity of CCs (Fageria, 2007; Meena et al., 2018). In some field trials, practicing legume CCs with high BNF rate (e.g., soybean and faba bean) has the potential to produce main-crop yields comparable with N fertilized treatments (McDonagh et al., 1995; Garba et al., 2022; Zhao et al., 2022). This indicates that environmental N pollution associated with fertilizer use can be possibly mitigated through

substituting chemical fertilizers with N-fixing green manure. However, detailed assessment of legume cover cropping impacts on global cropland remains a challenge, since the BNF rate in legumes is highly dependent on the effectiveness of rhizobial strains (Chen et al., 2016; Denton et al., 2017) and varies widely among sites and CC species (Liu et al., 2011; Ciampitti & Salvagiotti, 2018; Herridge et al., 2022).

1.4 Representing agricultural managements in Dynamic Global Vegetation Models

The benefits of soil nutrient and crop productivity from improved agricultural managements can be quantified at field sites and/or in controlled environments, but this empirical evidence is hard to scale up. Hence, this poses a challenge for large-scale modelling, for which dynamic global vegetation models (DGVMs) are useful due to their mathematical representation of vegetation and soil interactions under varying environmental and management conditions (Smith et al., 2014; McDermid et al., 2017; Pongratz et al., 2018; Herzfeld et al., 2021). Crops in DGVMs are parameterized as specific crop functional types (CFTs), which are groups of crops with similar agricultural plant traits (Bondeau et al., 2007; Levis et al., 2012; Drewniak et al., 2013). Progress over the last decade in representing cropland management strategies in DGVMs includes irrigation (Jägermeyr et al., 2015; Zhang et al., 2018), the implementation of dynamic sowing and harvest dates (Lindeskog et al., 2013), tillage (Ciais et al., 2011; Pugh et al., 2015; Lutz et al., 2019), residue management (Lindeskog et al., 2013; Ren et al., 2020), the incorporation of the N cycle enabling N fertilizer and manure application (Zaehle et al., 2011; Tian et al., 2012; Olin et al., 2015a; Von Bloh et al., 2018), and cover crops (Olin et al., 2015a; Porwollik et al., 2022).

Compared with site-level modelling studies, an assessment of the impacts of agricultural practices across regions or globally is still lacking, as a result of inadequate management information (e.g., spatial pattern of soil-conserving techniques) and missing or incomplete representation of land management options in models. For large-scale C-N cycle modelling assessments, alternative agricultural practices so far have been evaluated through stylized model setups with homogenous assumptions of management intensities (Olin et al., 2015a; Lutz et al., 2020; Jang et al., 2021). For example, Olin et al. (2015a) used the LPJ-GUESS DGVM to explore the impacts of soil-conserving strategies on SOC sequestration rate across global agricultural ecosystems, assuming that all cropland grid cells were to adopt no-till and 100% of residue retention managements. Similarly, to realistically reflect the spatial pattern of cover cropping, a recent modelling study performed by Porwollik et al. (2022) estimated with the LPJmL DGVM how conservation agriculture globally

11

might affect soil C-N and yields in response to non-legume CCs across four cropping systems. Their model results showed the potential of cover cropping for climate change mitigation via enhanced soil C pools, but the authors suggested that future modelling assessment of N-fixing CCs would be needed since this practice is identified as one practical strategy to address the conflict between the growing needs for crop production and the environmental problems associated with N fertilizer use (Abdalla et al., 2019; Nouri et al., 2022). To date, no study has applied DGVMs globally to investigate how N-fixing cover cropping affect agricultural ecosystem services, particularly in terms of cropland SOC sequestration, N losses, and crop yields.

1.5 Thesis structure and objectives

The main research questions of this thesis are:

- – How much nitrogen is fixed from the atmosphere in grain legumes globally?
- – Can legume cover crops contribute to environmental sustainability without compromising crop production in global croplands?
- – If so, how significant are these impacts from cover crops compared with other agricultural practices under present and future climate conditions?

To answer these questions, the results of this thesis are split into three main chapters (Chapters 3-5), each presenting the results of a peer-reviewed publication. Chapter 2 presents the common methodology used in all chapters of this thesis, as well as the specific methodologies used in Chapters 3-5. Chapter 3 describes how the incorporation of two new legume crops—soybean and faba bean—into the LPJ-GUESS DGVM improves the representation of agricultural production worldwide. The research questions addressed are:

- – What are the global spatial and temporal patterns of the nitrogen fixation rate in soybean and faba bean?
- – How much does the nitrogen fixed through BNF contribute to total nitrogen uptake in legumes?

Using the updated LPJ-GUESS model version from Chapter 3, Chapters 4 and 5 directly assess the large-scale influence of agricultural management practices on crop ecosystems, particularly with respect to soil carbon stocks, cropland nitrogen loss, and crop production. Chapter 4 focuses on potential effects from practicing two cover crop types—legumes and non-legumes—at a global scale. The research questions addressed are:

– To what extent can the implementation of cover cropping support carbon sequestration and soil nitrogen loss reduction?

– Do cover crop types, management duration, and nitrogen fertilization impact on the effectiveness of cover cropping in main-crop yields? If so, how?

In addition to cover crop managements assessed in Chapter 4, Chapter 5 comprehensively evaluates the possibility of seven management practices for achieving long-term environmental sustainability and food security. Chapter 5 focuses on eastern Africa as a case study, a region where agricultural soils have been experiencing strong degradation over recent decades. The research questions addressed in Chapter 5 are:

– Which alternative management practice is the best strategy for climate change mitigation via enhanced soil carbon pools?

– How could conservation agriculture contribute to crop ecosystems in the future?

Chapters 3-5 are based on the results from three papers: Chapter 3 builds upon the paper "Modeling symbiotic biological nitrogen fixation in grain legumes globally with LPJ-GUESS (v4.0, r10285)" by Ma et al. (2022) published in *Geoscientific Model Development*; Chapter 4 is based on the paper "Estimating the global influence of cover crops on ecosystem service indicators in croplands with the LPJ-GUESS model" by Ma et al. (2023) published in *Earth's Future*; Chapter 5 builds upon the paper "Assessing the impacts of agricultural managements on soil carbon stocks, nitrogen loss and crop production—a modelling study in eastern Africa" by Ma et al. (2022) published in *Biogeosciences*. The three published papers are attached in the Appendix of this thesis. A general conclusion and outlook section with respect to the overall research questions is presented in Chapter 6, providing a broader perspective on the findings of this thesis.

As much of the work presented in the following is involved with inputs from the co-authors in the published/submitted papers, I will use "we" instead of "I" throughout the Chapters 2-5.

2 Methods

2.1 Model description of LPJ-GUESS

The research tool used in this thesis is the Lund-Potsdam-Jena General Ecosystem Simulator (LPJ-GUESS). LPJ-GUESS is a process-based DGVM that can be used to investigate plant and soil C-N dynamics and their interactions in response to changes in environment (e.g., climate, atmospheric CO_2 levels, and N deposition) and management (e.g., crop type, N fertilizer, and harvest) through simulating individual- and patch-level plant physiological and biogeochemical processes on a daily time step (Smith et al., 2014). Natural vegetation implemented in the model is characterized by 12 plant functional types (PFTs), with ten woody and two herbaceous types included. PFTs differ in their phenology, photosynthetic pathway (C_3 or C_4), growth strategy, and bioclimatic limitations. Pastures are described as the competition between C_3 and C_4 grass PFTs, with half of aboveground biomass harvested annually to represent grazing impacts (Lindeskog et al., 2013). Four crop functional types (CFTs)—two temperate C_3 crops with spring and autumn sowing dates, a tropical C_3 crop representing rice, and a C_4 crop representing maize—are simulated to represent croplands, with crop-specific differences in morphological traits, dynamic C-N allocation patterns, heat requirements for growth, and N fertilization management (Olin et al., 2015b). For large-scale applications, the sowing date in each grid cell depends on a set of rules driven by crop- and climate-specific characteristics, with five seasonality types represented (see Waha et al. (2012) for details). Crops are prescribed as either rain-fed or irrigated and are harvested annually when the dynamic potential heat units (PHU; accumulated degree-days above a base temperature for each CFT) are fulfilled (Olin et al., 2015b). To account for post-harvest losses caused by mechanical damage or poor handling conditions, a harvest efficiency of 90% is used to adjust the modelled crop yields (Lindeskog et al., 2013). At present, within-year multi-cropping systems, which are common in tropical regions, have not been implemented in the model.

Cropland management options represented in LPJ-GUESS include irrigation, tillage, crop residue retention, N fertilizer and manure application, and cover crop grasses grown between two cropping seasons. Irrigation water is estimated as the amount of plant water deficit in the model and is added to the soil automatically when crops suffer from water stress. The effect of conventional tillage on heterotrophic respiration is simulated as a tillage factor of 1.94, which modifies the decay rate of four SOM carbon pools throughout the year and accelerates the soil decomposition on agricultural lands (Chatskikh et al., 2009; Pugh et al., 2015). In the standard LPJ-GUESS setup, 75% of

aboveground crop residue is removed from the fields after harvest; the rest, combined with root biomass, is assumed to enter to the soil litter pool for decomposition. Synthetic N fertilizer is added to the soil mineral N pool for plant uptake at three crop development stages, with varying application rates for each CFT. Manure is applied as a single input to cropland at sowing to account for the time required for manure N to be made available for crops. Manure is assumed to have a C:N value of 30 and is added to metabolic and structural SOM pools for decomposition (Olin et al., 2015a). Cover crops implemented in LPJ-GUESS so far have been simulated as competing temperate C_3 and tropical C_4 grasses grown annually between two consecutive growing seasons of main crops, replacing bare-soil fallow periods. Cover crop grass is sown on the 15th day after the harvest of the main crop, starting with a seedling that has an initial C mass of 0.01 kg C m^{-2} and C:N ratio of 16 (Olin et al., 2015a). Daily C and N mass in grasses are allocated to root and leaf pools based on a prescribed root:shoot partitioning ratio of 2 (Sainju et al., 2017), which is dynamically adjusted depending on plant water status. In the case of water stress, root allocation is increased (i.e., root:shoot partitioning ratio > 2) to help plants overcome the water limitation, following Penning de Vries et al. (1989). Cover crop grasses on fallow cropland in the simulations do not receive any management inputs (i.e., they grow under rain-fed and unfertilized conditions). Fifteenth days before planting the next main crop their shoot and root biomass are added to the surface litter and the soil metabolic/structural SOM pools, respectively, for further decomposition. Interplanting cover crops with main crops (i.e., two plants growing beside each other at the same time) is not implemented in the model.

C-N dynamics of the soils in LPJ-GUESS are modelled by 11 SOM pools differing in C:N ratios and resistance to decay, following the CENTURY model (Parton et al., 1993). Decomposition of SOM pools results in the release of CO_2 to the atmosphere (respiration) and C and N transfers between soil pools (Smith et al., 2014). C input to the receiver pool drives N mineralization or immobilization, as a result of maintaining mass balance and prescribed C:N ratios of the donor and receiver pool. Net N mineralization (i.e., mineralization minus immobilization), together with industrial N fertilizer and atmospheric N deposition, determine the size of the total soil mineral N pool, which is depleted by plant N uptake, as well as by crop ecosystem N losses through gaseous N emission and N leaching on a daily time step (Zaehle and Friend, 2010; Wårlind et al., 2014). Gaseous N emission produced in the soil to the atmosphere is simulated as NH_3, NO, N_2O and N_2, with the representation of soil N dynamic processes including ammonification, nitrification, and denitrification in the SOM pools. Following Parton et al. (1993), mineral N leaching in the model is

proportional to soil nitrate concentration and constrained by percolation rate and soil water content. N losses through soluble organic leaching are also added in LPJ-GUESS and determined by N decreases in soil microbial SOM nitogen pool (due to decomposition), water percolation, and soil sand fraction (Wårlind et al., 2014).

2.2 Implementation of legumes to LPJ-GUESS

Considering the importance of soybean in overall agriculture and trade, and the higher BNF rate in faba bean compared with other pulses (Peoples et al., 2021; Herridge et al., 2022), we incorporated these two grain legumes with BNF processes into LPJ-GUESS. In addition, a nitrogen-fixing grass—white clover (*Trifolium repens*), an herbaceous cover crop often used in conservation agriculture systems—is also included, by modifying the existing C_3 grass type (Olin et al., 2015a) but with BNF capabilities added (see Sect. 2.2.2 below). This updated LPJ-GUESS version (v4.0, r10285) will be used throughout the Chapters 3-5.

2.2.1 Updated daily carbon allocation parameters in grain legumes

Similar to most ecosystem and crop models, LPJ-GUESS uses accumulated heat requirements to simulate crop growth development (Lindeskog et al., 2013). To better represent C and N allocation in different phenological phases, Olin et al. (2015b) defined crop development stage by implementing the effects of temperature, verbalization days and photo-period, following Wang & Engel (1998). Here, we simplified the processes and assumed that the development of grain legumes is linearly correlated to its accumulated heat units, given the field-based soybean experiments (Irmak et al., 2013). It is estimated as:

$$DS = \begin{cases} a_{veg} + b_{veg} \times fphu & (fphu \leq fphu_{anthesis}) \\ a_{rep} + b_{rep} \times fphu & (fphu > fphu_{anthesis}) \end{cases} \tag{2.1}$$

where *DS* is crop development stage, ranging from 0 to 2 (*DS*=0, sowing; *DS*=1, flowering; *DS*=2, harvest); *fphu* is the fraction of today's accumulated heat units to total heat requirement; $fphu_{anthesis}$ is the threshold of *fphu* when anthesis starts, below (above) which crop growth belongs to the vegetative (reproductive) stage; and a and b are the linear regression coefficients, varying between the vegetative and reproductive phases. The values of a and b, and the crop-specific base temperature (°C) to estimate the accumulated heat units are both given in Table 2.1.

Allocation of assimilated C to leaves, stems, and roots is an important process before storage organs are formed. Unlike cereal crops, nodulated plants, particularly soybeans, are more likely to achieve

a higher photosynthesis rate and delay leaf senescence due to the continued N supply from N fixation (Abu-shakra et al., 1978; Kaschuk et al., 2010). A precise representation of assimilate partitioning to the plant organs when modelling BNF in grain legumes is especially important considering the high C cost from fixing N from the atmosphere. Productivity loss would be simulated if the leaf photosynthesis rate would not increase to compensate for the costs (Kaschuk et al., 2009).

Following Olin et al. (2015b), relationships between assimilate allocation to legume organs were established based on data from Penning de Vries et al. (1989) and Boote et al. (2002). We fitted the allocation functions using the Richards logistic growth curve (Eq. 2.2; Richards, 1959) to model the allocation to each organ dynamically and separately. Since the NPP cost to maintain BNF in the reproductive stage would reduce the flow of carbon assimilate to storage organs, we adjusted the allocation functions from Olin et al. (2015b) so that the model could dynamically adapt the allocation to grain over the seed-filling period in response to BNF cost (see Eqs. 2.3-2.5 for details).

$$f_i = a_i + \frac{b_i - a_i}{1 + e^{-c_i \times (DS - d_i)}} \tag{2.2}$$

where f_i represents the three allocation functions (see Eqs. 2.3-2.5 below); DS is crop development stage; a_i, b_i, c_i, d_i are fitting coefficients for the three functions, with the specific values given in Table 2.1.

2.2.1.1 Yield vs. the whole plant

After anthesis ($DS > 1$), most assimilates are allocated and retranslocated from the vegetative organs to the grains. During the late seed-filling period ($DS \geq d_1$, see Eq. 2.3), we assumed that the fraction of carbon allocated to yield would increase to partly compensate the productivity loss caused by spending on N fixation, with the cost of reducing the flow of carbon to leaves and stem (see Eq. 2.4). We established the ratio of the allocation to yield relative to the whole plant as:

$$f_1 = \frac{P_{yield}}{P_{veg} + P_{yield}} = \begin{cases} a_1 + \dfrac{b_1 - a_1}{1 + e^{-c_1 \times (DS - d_1)}} & DS < d_1 \\[3mm] \left(a_1 + \dfrac{b_1 - a_1}{1 + e^{-c_1 \times (DS - d_1)}} \right) \times (1 + P_{BNFcost}) & DS \geq d_1 \end{cases} \tag{2.3}$$

where P_{yield} and P_{veg} are the fraction of carbon allocated to yield and vegetative organs, respectively, ranging from 0 to 1; $P_{BNFcost}$ is the proportion of NPP used for BNF to today's total NPP; d_1 is the fitting coefficient, representing the DS of maximum growth rate of grain (see Table 2.1).

2.2.1.2 Leaf vs. shoot vegetative organs

Similarly, the ratio of leaf vs. shoot vegetative allocation is specified as:

$$f_2 = \frac{P_{leaf}}{P_{veg} - P_{root}} = \begin{cases} a_2 + \dfrac{b_2 - a_2}{1 + e^{-c_2 \times (DS - d_2)}} & DS < d_1 \\ \left(a_2 + \dfrac{b_2 - a_2}{1 + e^{-c_2 \times (DS - d_2)}}\right) - P_{BNFcost} & DS \geq d_1 \end{cases} \quad (2.4)$$

where P_{leaf} and P_{root} are the fraction of carbon allocated to leaf and root, respectively.

2.2.1.3 Root vs. vegetative organs

When a plant experiences water or nutrient stress, it invests more assimilate to roots relative to shoot vegetative organs (Penning de Vries et al., 1989). We implemented dynamic increases in the allocation to roots during the late seed-filling period to help legumes cope with the C loss from BNF cost, and established the relationship between the allocation to root and that to vegetative organs as:

$$f_3 = \frac{P_{root}}{P_{veg}} = \begin{cases} a_3 + \dfrac{b_3 - a_3}{1 + e^{-c_3 \times (DS - d_3)}} & DS < d_1 \\ \left(a_3 + \dfrac{b_3 - a_3}{1 + e^{-c_3 \times (DS - d_3)}}\right) + (1 - f_1) \times P_{BNFcost} & DS \geq d_1 \end{cases} \quad (2.5)$$

In addition, carbon partitioning to vegetative organs (P_{veg}) can be calculated by subtracting the reproductive allocation (i.e., P_{yield}) from the whole plant as:

$$P_{veg} + P_{yield} = 1 \Rightarrow P_{veg} = 1 - P_{yield} = 1 - f_1 \quad (2.6)$$

Finally, we can achieve dynamic carbon allocation to the plant organs over the growing season by combining Eqs. 2.3-2.6:

$$\begin{cases} P_{yield} = f_1 \\ P_{leaf} = f_2 \times (1 - f_1) \times (1 - f_3) \\ P_{stem} = P_{veg} - P_{root} - P_{leaf} = (1 - f_1) \times (1 - f_2) \times (1 - f_3) \\ P_{root} = f_3 \times (1 - f_1) \end{cases} \quad (2.7)$$

Table 2.1 List of parameters for the daily carbon allocation (Eqs. 2.1-2.7) for soybean and faba bean.

Parameter	Description	Soybean	Faba bean	Reference
a_{veg}	Regression intercept, vegetative phase	-0.06	-0.06	Irmak et al. (2013)
b_{veg}	Regression slope, vegetative phase	3.29	3.29	Irmak et al. (2013)
a_{rep}	Regression intercept, reproductive phase	0.71	0.71	Irmak et al. (2013)
b_{rep}	Regression slope, reproductive phase	1.31	1.31	Irmak et al. (2013)
$fphu_{anthesis}$	The threshold of fphu when anthesis starts	0.34	0.34	Irmak et al. (2013)
T_b	Base temperature for heat sum requirement	10 °C	4 °C	Etemadi et al. (2018)
f_1	a_1	0	0	This thesis
	b_1	1	1	
	c_1	8.93	9.59	
	d_1	1.41	1.46	
f_2	a_2	0.67	0.75	This thesis
	b_2	0	0	
	c_2	30.78	7.69	
	d_2	1.73	1.38	
f_3	a_3	0.56	0.59	This thesis
	b_3	0	0	
	c_3	3.74	5.53	
	d_3	0.53	0.51	

2.2.2 Representation of BNF

Fixing N from the atmosphere and N uptake from soils are two N sources for grain legumes to meet their total plant N demand. The latter has a higher priority for plants because the process needs less energy than N fixation (Macduff et al., 1996). Following on this idea, in LPJ-GUESS, N fixation is triggered when the following two assumptions are valid simultaneously (see Fig. 2.1): (a) today's plant growth still suffers from N limitation after N uptake from soils (i.e., the N deficit, plant N

demand minus soil N uptake, is greater than zero), and (b) today's NPP is positive, so that adequate C supply can be provided to fulfill the BNF cost.

The BNF scheme in LPJ-GUESS is adapted from previously published methods (Liu et al., 2011) in which it is parameterized as a combined response function to (a) the potential N fixation rate, (b) temperature, (c) soil water availability, and (d) the crop development stage:

$$N_{fix} = N_{fixpot} \times f_T \times f_W \times f_{DS} \tag{2.8}$$

where N_{fix} is the N fixation rate; N_{fixpot} is the potential N fixation rate; and f_T, f_W, f_{DS} are limitation functions of soil temperature, soil water status, and crop growth stage to BNF, respectively.

The potential N fixation rate is often related to the size and biomass of root nodules (Soussana et al., 2002; Voisin et al., 2003; Voisin et al., 2007). However, due to the difficulties in collecting roots and nodules under field conditions, some studies adopted aboveground biomass as an alternative to estimate the legume BNF rate, following the empirical relationship between shoots and roots (Yu et al., 2002; Corre-Hellou et al., 2009; Wu et al., 2020). Given the absence of the nodulation process in LPJ-GUESS at present, N_{fixpot} is assumed to be linearly related to root dry matter:

$$N_{fixpot} = N_{maxfixpot} \times DM_{root} \tag{2.9}$$

where $N_{maxfixpot}$ is the maximum nitrogen fixation rate of roots (g N g^{-1} root DM) and DM_{root} is root dry matter (g root DM m^{-2}). As the experiment-based parameter $N_{maxfixpot}$ is highly dependent on the effectiveness of rhizobial strains, and shows considerable variation between species and sites, it is difficult to obtain this parameter for large-scale applications. Here, we assume that legume plants are inoculated (or there are high enough populations of strains in the soil) and grown in a robust soil ecosystem, so that $N_{maxfixpot}$ is not constrained by the absence of rhizobia. For the two grain legumes, $N_{maxfixpot}$ is assumed to be a constant of 0.03 g N g^{-1} root DM, a moderate value taken from the literature (Eckersten et al., 2006; Boote et al., 2009). For white clover grass, $N_{maxfixpot}$ is set to 0.012 g N g^{-1} root DM, a mean value from the field-based range of 0-0.03 g N g^{-1} root DM reported in Michaelson-Yeates et al. (1998) and Soussana et al. (2002).

Soil temperature is a major factor affecting both microbial activities and plant growth. For soybean, the optimal soil temperature for N fixation can range between 20-35°C (Boote et al., 2009). For white clover, Halliday & Pate (1976) observed in experiments that nitrogenase activity appeared to have a broad optimum range of 13-26°C with a sharp decline below 13°C and above 26°C. The

influence of soil temperature on legume BNF is represented in the model as a number of linear relationships:

$$f_T = \begin{cases} 0 & (T < \text{T}_{min} \ or \ T > \text{T}_{max}) \\ \frac{T - T_{min}}{T_{optL} - T_{min}} & (\text{T}_{min} \leq T < \text{T}_{optL}) \\ 1 & (\text{T}_{optL} \leq T \leq \text{T}_{optH}) \\ \frac{T_{max} - T}{T_{max} - T_{optH}} & (\text{T}_{optH} < T \leq \text{T}_{max}) \end{cases}$$
(2.10)

where T is soil temperature (°C) at 25 cm depth, representing the mean temperature of the topsoil layer in the model (0-50cm); T_{min} (T_{max}) is the minimum (maximum) temperature below (above) which N fixation stops; and T_{optL} and T_{optH} are the lower and higher optimal temperatures within which N fixation is not limited by temperature. The values of these four temperature thresholds vary among legume plants and are shown in Table 2.2.

Water deficit as well as waterlogging can also affect N fixation. Too little or too much water dramatically inhibits BNF due to impacts of drought stress and oxygen deficit, respectively, on nodule nitrogenase activity (Marino et al., 2007). Following previously published methods (e.g., APSIM, Robertson et al., 2002; EPIC, Cabelguenne et al., 1999; SOILN, Wu & McGechan, 1999), a linear water-limitation function is incorporated into LPJ-GUESS, and is represented as:

$$f_W = \begin{cases} 0 & (W_f \leq W_a) \\ \varphi_1 + \varphi_2 \times W_f & (W_a < W_f < W_b) \\ 1 & (W_f \geq W_b) \end{cases}$$
(2.11)

where W_f is relative soil water content in the top soil layer (0-50cm), ranging between 0 and 1; φ_1 and φ_2 are empirical coefficients; W_a is the threshold of W_f below which N fixation is fully restricted by soil water deficit and W_b is the value above which N fixation is not inhibited by soil water content. The values of the parameters are given in Table 2.2.

Much experimental evidence has indicated that the N fixed by grain legumes varies widely among growth stages, reflecting the development of root nodules from establishment to senescence (Córdova et al. 2020). In the CROPGRO model (Boote et al., 2009), N fixation in soybean starts in the early vegetative stage and continues until the end of physiological maturity, whereas it ceases at the middle of the seed-filling period in the EPIC model (Cabelguenne et al., 1999). In this study, a more specific function, similar to the temperature response function, is implemented to the BNF scheme to represent the variation of N fixation with the course of life cycle in grain legumes:

$$f_{DS} = \begin{cases} 0 & (NDS < \text{NDS}_{\text{min}} \text{ or } NDS > \text{NDS}_{\text{max}}) \\ \dfrac{NDS - NDS_{min}}{\text{NDS}_{\text{optL}} - \text{NDS}_{\text{min}}} & (\text{NDS}_{\text{min}} \leq NDS < \text{NDS}_{\text{optL}}) \\ 1 & (\text{NDS}_{\text{optL}} \leq NDS \leq \text{NDS}_{\text{optH}}) \\ \dfrac{\text{NDS}_{\text{max}} - NDS}{\text{NDS}_{\text{max}} - \text{NDS}_{\text{optH}}} & (\text{NDS}_{\text{optH}} < NDS \leq \text{NDS}_{\text{max}}) \end{cases} \qquad (2.12)$$

In contrast to grain legumes, white clover, as an herbaceous perennial plant in cover cropping systems, is often killed chemically or mechanically after several months of growth to prevent potential penalties to the subsequent main crops. Nodule senescence is thus not a major factor affecting N fixation in grass-based cover cropping. Accordingly, we modify Eq. 2.12 to only account for the nodule establishment effects on N fixation at the early development stage of white clover:

$$f_{DS} = \begin{cases} 0 & (NDS < \text{NDS}_{\text{min}}) \\ \dfrac{NDS - NDS_{min}}{\text{NDS}_{\text{optL}} - \text{NDS}_{\text{min}}} & (\text{NDS}_{\text{min}} \leq NDS < \text{NDS}_{\text{opt}}) \\ 1 & (NDS \geq \text{NDS}_{\text{opt}}) \end{cases} \qquad (2.13)$$

where NDS is normalized legume development stage, ranging from 0 to 1; NDS_{min} is the time before which there is no N fixation due to inadequate nodulation; NDS_{max} is the time after which N fixation suspends due to nodule senescence; and NDS_{optL} and NDS_{optH} define the period within which legume BNF rate is not inhibited by development stage. The values of the parameters for soybean, faba bean, and white clover are derived from the literature and listed in Table 2.2.

Apart from the environmental limitation factors, soil mineral N concentration and the amount of daily NPP also affect BNF rate. The NPP requirement for BNF costs (Eq. 2.14 below) in LPJ-GUESS is computed based on the estimated N fixation rate (N_{fix}; Eq. 2.8) by multiplying the C cost per unit fixed N, which is set to a constant of 6 g C g^{-1} N fixed as a moderate value taken from previous studies (Patterson & Larue, 1983; Boote et al., 2009). Kaschuk et al. (2009) reported that fixing N from the atmosphere would cost 8-32% of net photosynthetic C to maintain legume symbiotic growth, activity, and reserves. To catch rare cases where the simulated C cost may exceed the reported range, we thus assume that at most 50% of today's NPP in legumes can be used for N fixation. When daily C cost is more than 50% of NPP, the modelled BNF rate (N_{fix}; Eq. 2.8) is further adjusted as:

$$P_{BNFcost} = 6 \times N_{fix} / NPP_{today} \qquad (2.14)$$

$$N_{fix\_today} = \begin{cases} N_{fix} & P_{BNFcost} < 0.5 \\ 0.5 \times NPP_{today}/6 & P_{BNFcost} \geq 0.5 \end{cases} \tag{2.15}$$

where $P_{BNFcost}$ is the fraction of today's NPP used for BNF; N_{fix} is the estimated BNF rate from Eq. 2.8; NPP_{today} is today's NPP; N_{fix_today} is today's final fixed N in legumes. The fixed N is partly transported to plant leaves and continues to support the photosynthesis activity next day, resulting in additional C benefits by reducing N limitation on leaf carboxylation capacity. More details on the BNF scheme in LPJ-GUESS can be found in Fig. 2.1.
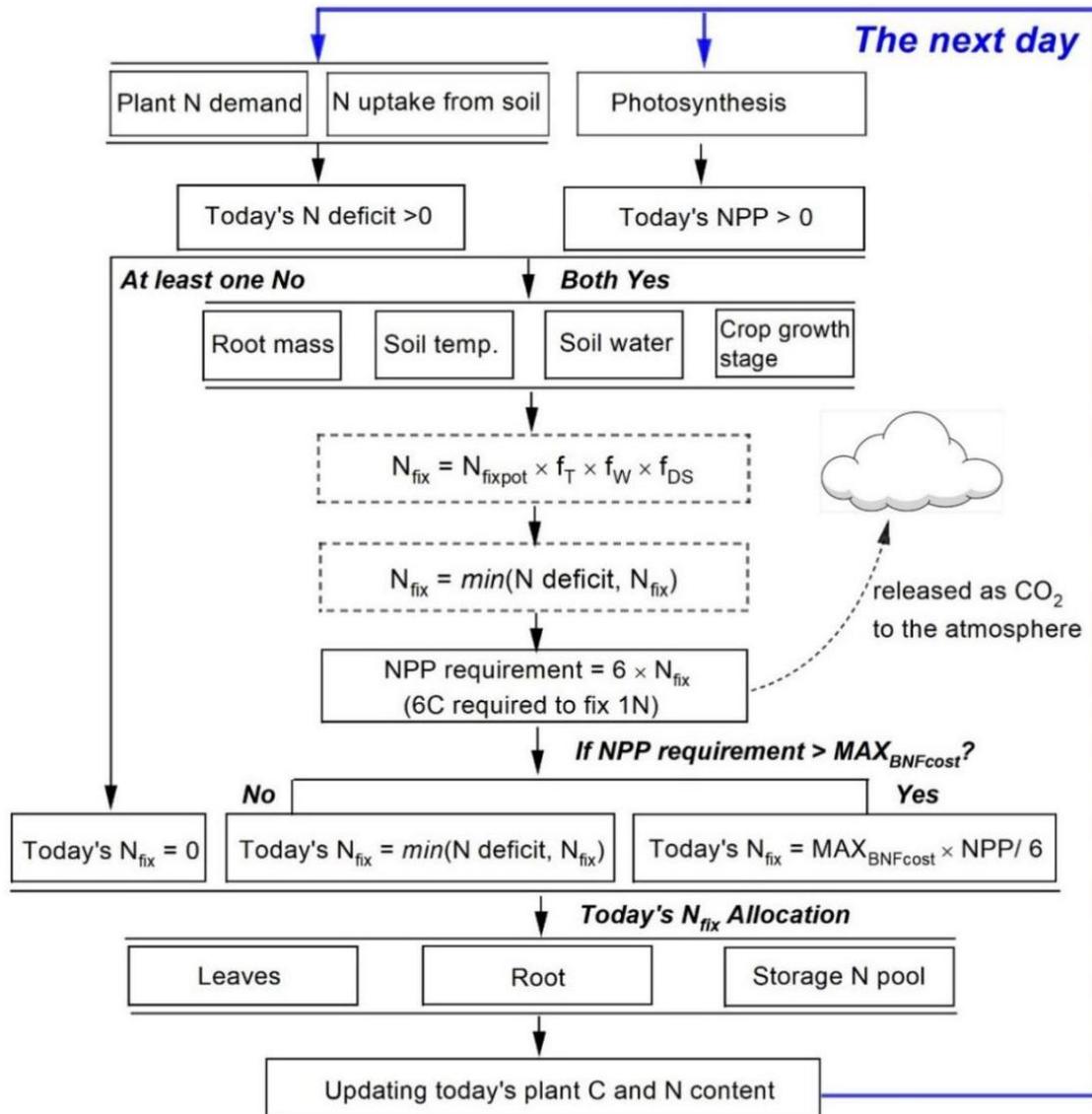
Figure 2.1 Representation of the N fixation route used in grain legumes in LPJ-GUESS. Today's N deficit is calculated as the difference between plant N demand and soil mineral N uptake. N_{fix} in dotted boxes are intermediate values.

Table 2.2 Overview of BNF-related variables and parameters used in the model for three legumes.

Parameter	Description	Soybean	Faba	White clover	Unit
N deficit	plant N demand minus soil N uptake	dynamic	dynamic	dynamic	g N m^{-2} d^{-1}
NPP	net primary productivity	dynamic	dynamic	dynamic	g C m^{-2} d^{-1}
$N_{maxfixpot}$	maximum BNF rate of roots	0.03	0.03	0.012	g N g^{-1} root DM
DM_{root}	root dry matter	dynamic	dynamic	dynamic	g root DM m^{-2}
C cost	carbon cost per unit fixed N	6	6	6	g C g^{-1} N fixed
T	soil temperature at 25 cm depth	dynamic	dynamic	dynamic	°C
T_{min}	the minimum temperature for the start of BNF	5	1	9	°C
T_{optL}	lower bound of optimal temperature for BNF	20	16	13	°C
T_{optH}	upper bound of optimal temperature for BNF	35	25	26	°C
T_{max}	the maximum temperature for the stop of BNF	44	40	30	°C
W_f	relative soil water content (0-50 cm)	dynamic	dynamic	dynamic	fraction
W_a	lower bound of water content below which BNF is limited by soil water	0.2	0	0	fraction
W_b	upper bound of water content above which BNF is not limited by soil water	0.8	0.5	0.5	fraction
φ_1	coefficient of soil water content	-0.33	0	0	–
φ_2	coefficient of soil water content	1.67	2	2	–
NDS	normalized crop development stage	dynamic	dynamic	dynamic	–
NDS_{min}	the minimum development stage for the start of BNF	0.1	0.1	0.1	–
NDS_{optL}	lower bound of development stage for BNF	0.3	0.3	0.3	–
NDS_{optH}	upper bound of development stage for BNF	0.7	0.6	–	–
NDS_{max}	the maximum development stage for the stop of BNF	0.9	0.8	–	–

2.3 Specific methodology in Chapter 3

In Chapter 3, field-based soybean and faba bean data from published sources, together with global yield statistics from legume-producing countries and region-level N fixation data from the literature, are compared with LPJ-GUESS model runs to examine the performance of the new implementation in simulating BNF rates and yields from site scale to a large region.

In order to build up the stabilized soil C-N levels on cropland, all LPJ-GUESS simulations are initialized with a 500-year spin-up using atmospheric CO_2 concentration from 1901 and repeating de-trended climate from 1901-1930 (see Sect. 2.3.1 for data information). During spin-up, potential natural vegetation (PNV) is simulated for the first 470 years, and then the cropland fraction linearly increases from zero to the first historic value (1901) in the last 30 years. This protocol of model spin-up will be used throughout the Chapters 3-5.

2.3.1 Model evaluation at site scale

To evaluate the model's ability to simulate BNF rates and yields, N fixation trials under field conditions with detailed measurements of soil N uptake and biomass were collected from the literature. A total of 17 soybean and 7 faba bean sites between ~33°S and ~53°N were compiled (Fig. 2.2). In this data set, grain yield, dry biomass and N mass of plant various tissues, together with the proportion of plant N derived from the atmosphere (%Ndfa), soil N uptake and N fixation were widely-reported variables. These data were thus chosen as target variables for model evaluation. Additionally, to convert plant C mass to dry matter, a conversion factor of 2.0 was used. Dry weight was converted to wet weight by assuming a water fraction of 0.13 in grain legumes (Córdova et al., 2019).

Since specific leaf area (SLA) and target grain C:N ratio play a significant role in determining N uptake and N retranslocation to grain during seed-filling in the model (Camargo-Alvarez et al., 2022), we conducted two simulations to explicitly examine model performance across all sites. For 'site-specific' runs, the reported SLA and grain C:N ratio were used for the simulation at sites for which these were available. For 'global-uniform' parameter runs, SLA was represented as a constant of 40 m^2 kg^{-1} C for soybean and 45 m^2 kg^{-1} C for faba bean (Penning de Vries et al., 1989), and target grain C:N ratio was set to 8 and 10, respectively (Kattge et al., 2020).

Because weather data for most evaluated sites was not available, a gridded daily climate data set at 0.5° resolution from GSWP3-W5E5 (Lange, 2019; Cucchi et al., 2020) was used as input (air

temperature, precipitation, and solar radiation), choosing the grid cell where the experimental sites were located. Likewise, there was not much information on land use during the years preceding the field trials for most sites. Therefore, to maintain SOM pools in equilibrium after model spin-up, we decided to implement a common cropping system of maize-legume rotation annually from 1901 to the year before the trials start, with no N fertilizer applied to both crops. Over the trial period, the management practices were implemented based on information provided in the literature. In addition, site-specific soil physical properties—bulk density, fractions of sand, silt, and clay—derived from the literature were used as external forcing to further calculate corresponding soil water characteristics in the model.

The agreement between modelled and observed variables was assessed using adjusted R^2 (the goodness of fit for the linear regression analysis), mean error (ME), mean absolute error (MAE), root mean square error (RMSE), and Pearson correlation coefficient (r):

$$ME = \frac{1}{n}\sum_{i=1}^{n}(\frac{M_i - O_i}{O_i}) \times 100\% \tag{2.16}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(\frac{|M_i - O_i|}{O_i}) \times 100\% \tag{2.17}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(M_i - O_i)^2} \tag{2.18}$$

$$r = \frac{\sum_{i=1}^{n}(M_i - \bar{M})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^{n}(M_i - \bar{M})^2 \sum_{i=1}^{n}(O_i - \bar{O})^2}} \tag{2.19}$$

where M_i and O_i indicate modelled and observed values, with the mean value given as $\bar{M}$ and $\bar{O}$ respectively. n is the number of observations.
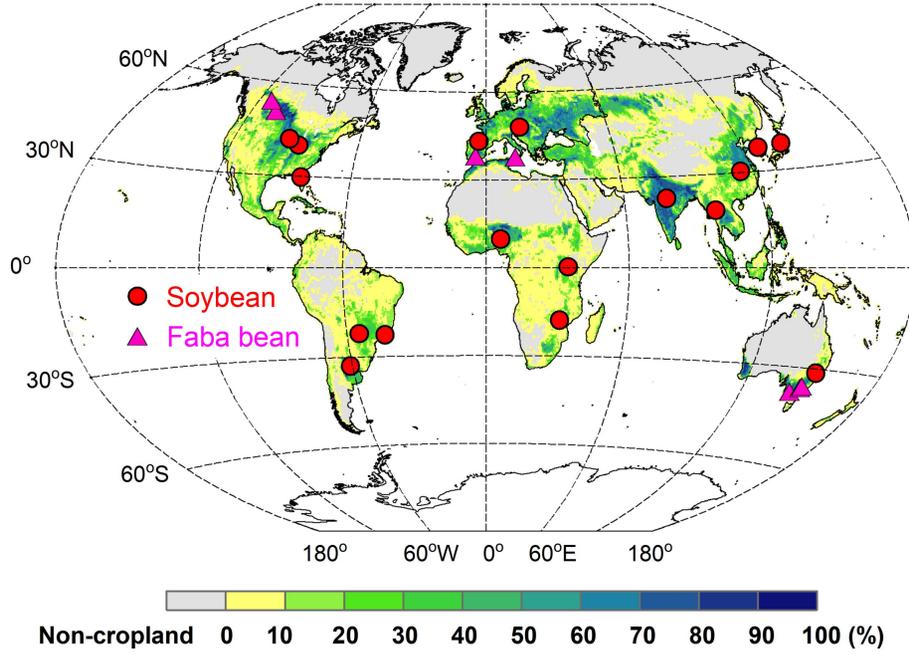
Figure 2.2 Soybean (red circles) and faba bean (magenta triangles) sites used for BNF evaluation. The map background is cropland fraction (%) averaged over 1996-2005 at the resolution of 0.5° × 0.5°, derived from the LUH2 data set (Hurtt et al., 2020).

2.3.2 Global yields and BNF rates

To test model performance in simulating large-scale legume yields and BNF rates, FAO-reported national yields (FAOSTAT, 2023) were collected and compared with LPJ-GUESS output. Furthermore, Peoples et al. (2009) collated the N fixation rates in widely-grown legumes from a range of published sources and divided them into different geographical regions. In order to compare our simulated BNF with the literature reports, each simulated 0.5°×0.5° grid cell was classified to be in one of the following ten regions given in Peoples et al. (2009): West Asia, South Asia, South-East Asia, East Asia, Central Asia, Africa, Europe, North America, South America, and Oceania.

For regional comparison, the modelled gridded yield and BNF rate were aggregated to national and continental scales, respectively, using information of crop-specific cover area on spatial pattern:

$$Var_{region} = \frac{\sum_{i=1}^{n} [(Var_{rain})_i \times (Area_{rain})_i + (Var_{irri})_i \times (Area_{irri})_i]}{\sum_{i=1}^{n} [(Area_{rain})_i + (Area_{irri})_i]} \tag{2.20}$$

where *Var* is yield or BNF rate; Var_{region} is the aggregated result in a given region; *i* is the grid cell number in that region, ranging from 1 to n; Var_{rain} and Var_{irri} represent the modelled yield or BNF

rate under rain-fed and irrigated conditions, respectively; $Area_{rain}$ and $Area_{irri}$ are the crop-specific rain-fed and irrigated areas used in simulations, respectively (see below).

Similar to the site-level setup in Sect. 2.3.1, the daily climate data set GSWP3-W5E5 was also used for the global simulation, spanning from 1901-2016 at 0.5° resolution. Annual atmospheric CO_2 concentration was taken from Meinshausen et al. (2020). Historical land use/land cover input data between 1901 and 2014 were adopted from LUH2 (Hurtt et al., 2020) and remapped from 0.25° to 0.5° with fractions of natural vegetation, pasture, and cropland given for each grid cell. The growth distribution of various crop types, distinguishing shares of rain-fed and irrigated crop-specific fraction per grid cell, was based on the MIRCA data set around the year 2000 (Portmann et al., 2010). Since no detailed information was available on the fractional cover of faba bean, the 'pulse' fraction in MIRCA was used as input instead, and 'pulses' country-level yield statistics provided by FAOSTAT (2023) were collected to compare with faba bean outputs by LPJ-GUESS. In addition, to parameterize soil hydraulic properties, cropland soil texture classes in the upper soil layer (0-30 cm) from ISIMIP/GGCMI phase 3 (Volkholz & Müller, 2020) were used and held constant over the course of the model experiments.

In terms of timing of N fertilizer application, a recent synthetic analysis by Mourtzinis et al. (2018) showed that splitting N application at planting and early reproductive stage led to dramatically higher soybean yields than a single application method. Mineral N fertilizer was thus applied to legumes at the time of sowing and flowering with equal applications in our implementation. To realistically account for the time required for manure N to be available to plants, manure was added to soils at the time of sowing as a single application. Crop-specific mineral N fertilizer and manure inputs from 1901-2014 were taken from Ag-GRID (Elliott et al., 2015) and Zhang et al. (2017), respectively. Moreover, monthly atmospheric N deposition simulated by CCMI from 1901 to 2014 was used and interpolated to the same resolution of the climate forcing (0.5°×0.5°) (Tian et al., 2018).

2.4 Specific methodology in Chapter 4

This chapter is divided into two parts. First, we test the model's ability to reproduce the observed responses in SOC stocks, N leaching, and crop yields to N-fixing and non-N-fixing cover crops (CCs) at various field trial sites around the world. Next, we perform four global simulations of CC cultivation in different tillage systems (Table 2.3). Our analyses focus on the earlier mentioned three ecosystem service indicators, first evaluating the model results against estimates from global-

level studies and statistics, then analyzing and discussing the potential contribution of CCs to environmental sustainability and food security under three conservation agriculture scenarios (see Sect. 2.4.2 for details). Model simulation spin-up follows the protocol run described in Chapter 3 (see Sect. 2.3 above). All input data used for simulations are summarized in Table 2.4, with the specific experimental setups explained in detail below.

2.4.1 Model evaluation at site scale

To examine the model performance, cover crop field trials that also report observations of SOC stocks, N leaching, and crop yields were collected from the existing literature. Due to the absence of intercropping systems in the model, we only selected field trials in which CCs were either grown during the bare fallow period or undersown in main crops. For the latter case, CCs usually coexist with the main food crops for a short while (ca. 1-2 months before the main crop is harvested); CC growth is dormant during the winter months, but continues in spring, and CC crops are then terminated several days prior to the next planting of the main crop (Valkama et al., 2015). Additionally, to capture the variability of the observed data, CC treatments needed to cover at least two growing seasons, with the whole plant used as green manure or mulch returning to the fields. As a result, a total of 43 studies carried out at 41 different sites were compiled for evaluation. These studies investigated the effects of two cover crop functional types—legumes (CC_L) and non-legumes (CC_{NL})—on soil C sequestration (12 sites), N leaching (13 sites), and crop yields (29 sites) across four cropping systems (wheat, maize, rice, and soybean) and under various water and N management practices (Fig. 2.3).

Following the site-level simulation setup in Chapter 3 (Sect. 2.3.1), the gridded climate data set GSWP3-W5E5 at 0.5° resolution was used as input due to the absence of weather data for most selected sites. Also, because there was not much information on land use during the years preceding the field trials, we assumed that all sites were under grassland systems from 1901-1905, followed by a cropland period of 1906-1910, with this 5-year alternation between grassland and cropland repeated until the field trials began. Since cropland at most sites had already been present for several years at the beginning of the CC experiment, we simulated five years of cropland preceding the site trials at those locations for which no other information was reported. Over the experimental period, model runs were performed according to management information reported in the literature. At the moment LPJ-GUESS does not simulate the cultivation of two crops simultaneously on the same field, whereas undersown CCs in the field experiments are generally grown together with

main crops at least one or two months. To better represent the total length of the cover crop growing season in the model simulations, we adjusted the sowing date of undersown CCs (referred to as the 4-A1 runs in Table 2.4) to one day after the main crop harvest (instead of the default 15) and terminated the plants one day before the establishment of the next primary crop. For CCs solely grown on fallow cropland (4-A2 runs; Table 2.4), their planting and harvest dates were assumed to be the same as the LPJ-GUESS standard setup, following the common field practice at most sites (Mazzoncini et al., 2011; Kaspar et al., 2012; Duval et al., 2016).
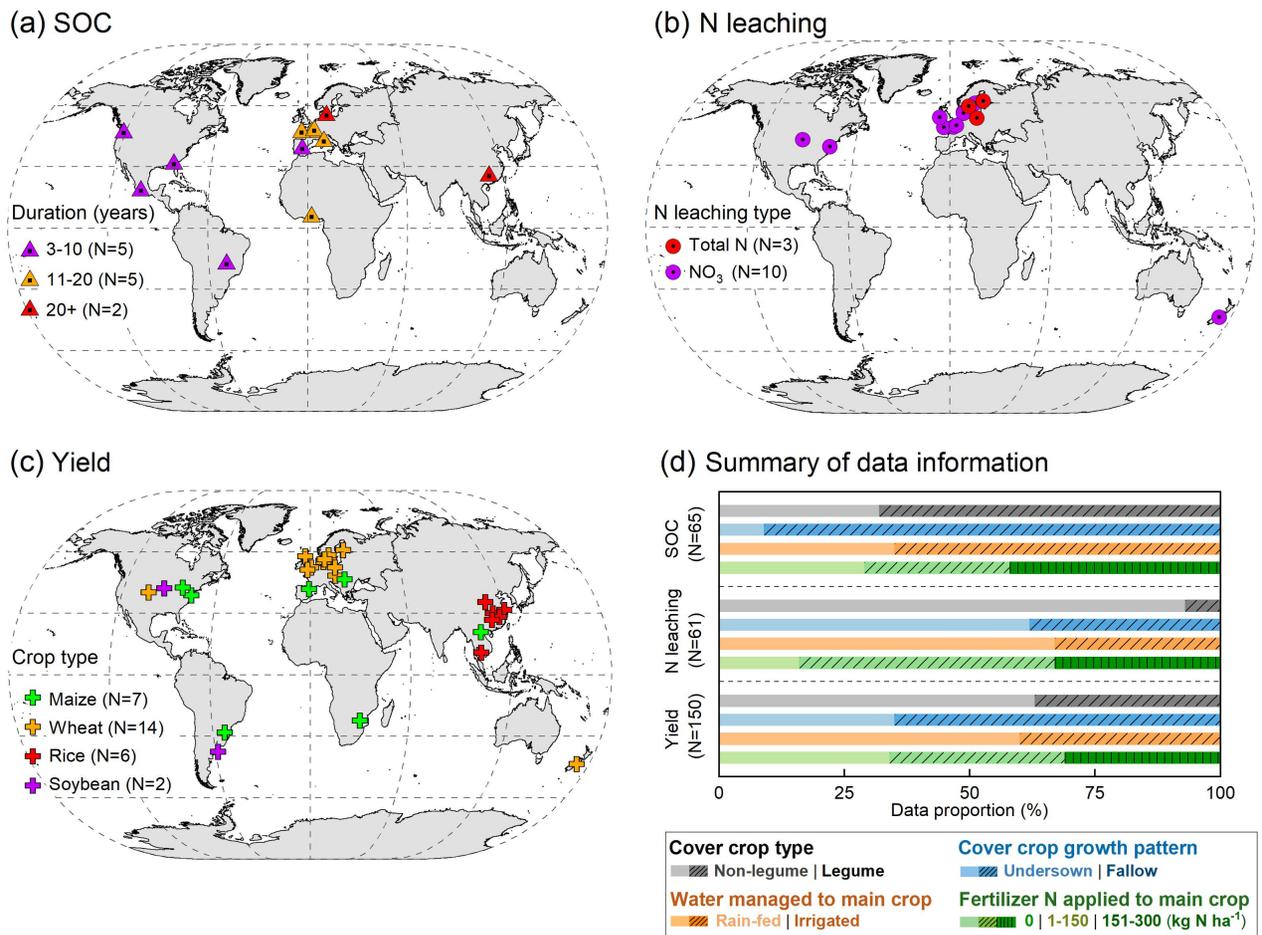
Figure 2.3 Distribution of cover cropping field studies used for model evaluation of cropland SOC stocks (a), N leaching loss (b), and crop yield (c). All studied SOC sites (12) had continuously practiced CC cultivation for more than three years, and the leached N loss at the evaluated sites (13) were reported as either total N (mineral plus organic) or nitrate (NO_3). The influence of CC practice on crop production was investigated in four cropping systems (maize, wheat, rice, and soybean) at 29 sites from 16 countries. Overall measurement information in cover crop field experiments—cover crop types (legumes or non-legumes), growth patterns (undersown or fallow), and water and N fertilizer managements to main crops—is shown in (d).

2.4.2 Global agricultural ecosystem response to cover cropping

In this experiment we performed simulations with four CFTs—wheat (including spring wheat and winter wheat), maize, rice, and soybean—which jointly provide more than two-third of the world's food supply (FAOSTAT, 2023). To detect how CCs affect cropland ecosystem services, two cover crop types—leguminous (CC_L) and non-leguminous (CC_{NL}) grasses—were assessed. An additional combined practice, with N-fixing cover crop and no tillage (CC_LNT), was used to represent important aspects of conservation agriculture. Model outputs of these three practices were compared to a control simulation with bare fallow (NoCC), applying the simulation setup given in Table 2.3.

The model experiments started with a baseline simulation of the historical period (1901-2014) under NoCC management after model spin-up, using dynamic gridded climate, land use/land cover, and N fertilizer data ($0.5°\times0.5°$), together with atmospheric CO_2 concentration. The result of this run was to produce present-day SOM pools on off-season fallow cropland across the globe (Table 2.4). This baseline simulation is referred to as 4-B1. Subsequent runs of four management practices listed in Table 2.3 branched from this present-day state in 2015 and are referred to as the 4-B2 runs. These simulations ran for 36 years (the maximum duration found in cover cropping field trials in our analyzed sites) but are not intended to estimate SOC storage, N leaching and crop production through 2050; rather, they are designed to detect the relative changes in these three ecosystem indicators due to replacing bare fallows with cover crops. For that reason, we use constant repeated 1995-2014 climate with temperature de-trended, combined with 2014 land use, fertilizer, manure, and CO_2 concentration (Table 2.4). In order to contrast short- with long-term cover crop impacts, model outputs in the first (years 1-10) and last (years 27-36) decades were used for analysis.

Rather than the daily weather data that can be used for spin-up and historical simulations, during such "extended" simulations, LPJ-GUESS can only be driven by monthly de-trended climate. The observation-based monthly data set CRUJRA v2.1 was thus adopted for all simulation described in this subsection. Specifically, we used mean air temperature, total precipitation, solar radiation, and number of wet days from 1901-2014 at $0.5°$ resolution (Kobayashi et al., 2015; Harris et al., 2020). The rest of the model input data—including atmospheric CO_2 concentration, historical land use/land cover, CFTs-specific growth distribution, synthetic N fertilizer and manure application, N deposition, and soil physical properties—were consistent with global simulation setups in Chapter 3 (see Sect. 2.3.2). Since large-scale statistics on actual cover crop acreage do not exist, the conservation agriculture (CA) area was used to represent the potential cover crop distribution on

croplands, following setups in a recent modelling study (Porwollik et al., 2022). We performed all global simulations under three CA area scenarios: (1) CA_{his}, representing the approximate area of CA practice currently adopted in global croplands; (2) CA_{pot}, representing the potential agricultural lands that might implement CA systems under present socio-economic and soil biophysical conditions; (3) CA_{all}, assuming all cropland that was under CA management. Spatial pattern of CA_{his} and CA_{pot} were taken from a gridded data set developed by Porwollik et al. (2019), in which national FAO-reported CA area around the year 2005 was downscaled to grid cell level and the potential CA-suitable agricultural lands were estimated based on a range of rule-based approaches. To characterize the CA_{all} scenario, LUH2 land use data at the year 2014 were used (Hurtt et al., 2020). The spatial distribution of these three CA scenarios, as well as their total areas, are shown in Fig. 2.4.
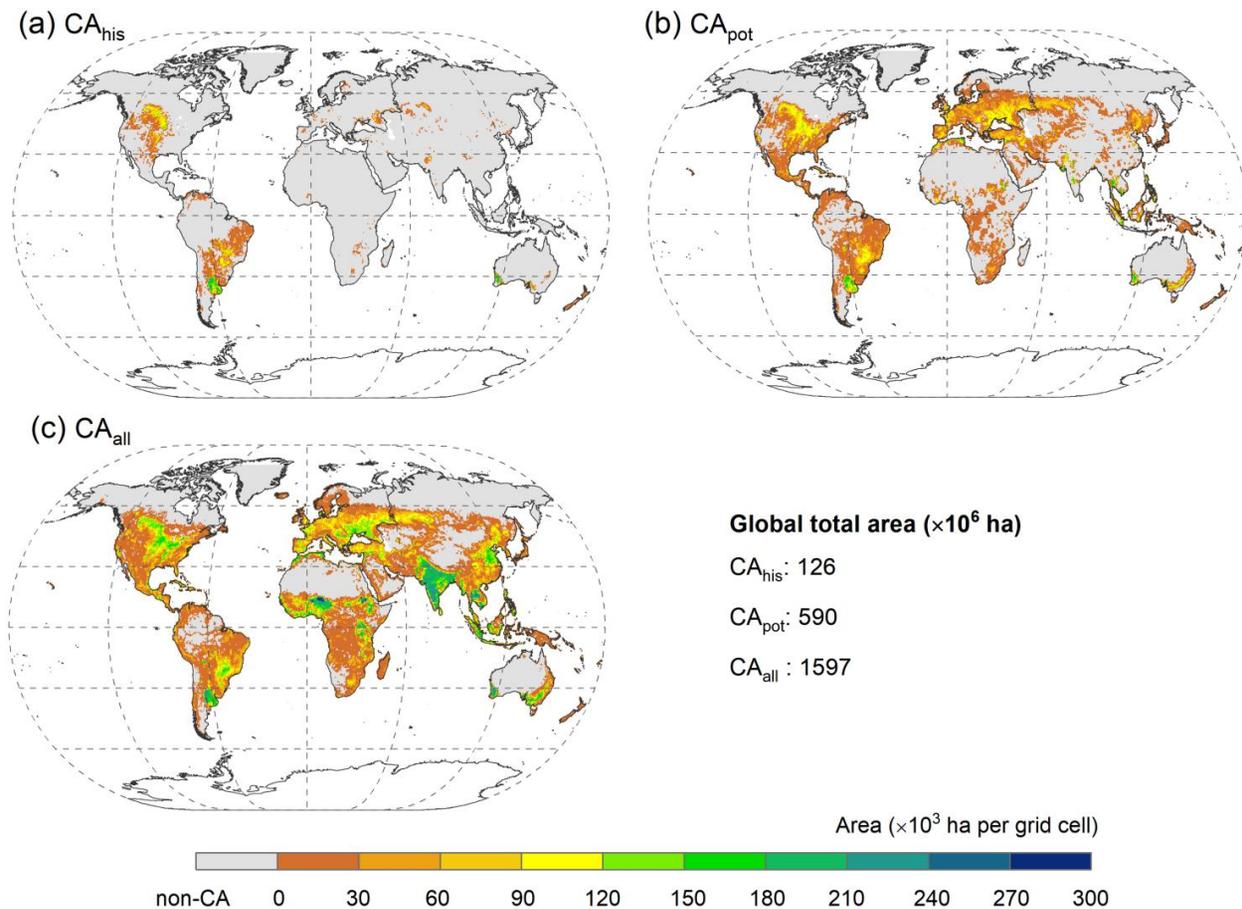
Figure 2.4 Maps of three conservation agriculture (CA) area scenarios for cover crop managements assumed in this study: (a) current CA areas adopted in global cropland (CA_{his}); (b) potential agricultural lands that might implement CA practices (CA_{pot}); (c) all cropland under CA managements (CA_{all}).

Table 2.3 Global-scale simulation setups representing different cover crop managements in Chapter 4.

Simulation*	NoCC	CC$_L$	CC$_{NL}$	CC$_L$NT
Legume cover crop	No	N-fixing C$_3$ grass	No	N-fixing C$_3$ grass
Non-legume cover crop	No	No	Competing C$_3$ and C$_4$ grasses	No
Residue retention	25%	25%	25%	25%
Manure application	Yes	Yes	Yes	Yes
Mineral N fertilizer	Yes	Yes	Yes	Yes
Tillage	Yes	Yes	Yes	No

* Abbreviations: NoCC – control treatment with bare fallows; CC$_L$ – legume cover crops; CC$_{NL}$ – non-legume cover crops; CC$_L$NT – combined management practice with legume cover crops and no tillage.

2.4.3 Data analysis

Model performance at site scale was evaluated by comparing the simulated and observed ecosystem service indicators—SOC stocks, N leaching loss, and crop yield—in response to the implementation of cover crops. For SOC stocks comparison, when the observed values in some field experiments were only provided as concentrations (g kg^{-1}), we converted these to stocks (Mg ha^{-1}) using the following equation:

$$SOC_{stock} = (SOC_{con} \times BD \times D)/10 \tag{2.21}$$

where SOC_{stock} and SOC_{con} represent soil organic carbon stocks (Mg ha^{-1}) and concentration (g kg^{-1}), respectively. BD is bulk density (g cm^{-3}) and D is soil depth (cm).

The sampled soil depth for SOC and N leaching in our compiled data set varied from 15-40 cm and 60-150 cm, respectively. To compare model outputs with observations, we standardized the measured SOC and N leaching from the original depth to the modelled depth of 150 cm, following the depth distribution function developed by Jobbágy & Jackson (2000) and further described by McClelland et al. (2021):

$$Y = 1 - \beta^D \tag{2.22}$$

$$VAR_{150} = \frac{1 - \beta^{150}}{1 - \beta^{D0}} \times VAR_{D0} \tag{2.23}$$

where Y is the cumulative proportion of the SOC or N leaching from the surface to depth D (cm) and β is the relative rate of decrease in these two variables with soil depth. The value of β was obtained from a meta-analysis study and set to 0.9786 for SOC and 0.9831 for N leaching (Abdalla et al., 2019). VAR denotes SOC or N leaching; D_0 is the original soil depth available in the literature; VAR_{150} and VAR_{D0} represent the cumulative SOC stocks or N leaching at 0-150 cm and original soil depth, respectively.

Based on these post-processed site-level observed data, the accuracy of the model in predicting cropland SOC stocks, N leaching, and crop yield was assessed using adjusted R^2, ME, MAE, and RMSE given in Sect. 2.3.1. In addition, to quantify the response of cropland soil C storage to CCs in comparison with the control treatment (NoCC), the annual SOC sequestration rate, ΔSOC_{rate} (Mg C ha^{-1} yr^{-1}), was calculated as:

$$\Delta SOC_{rate} = \frac{SOC_X - SOC_{NoCC}}{YR} \tag{2.24}$$

where SOC_X and SOC_{NoCC} are the respective SOC stocks under the cover crop and control treatments, x denotes any cover crop practices (CC$_L$, CC$_{NL}$, and CC$_L$NT; see Table 2.3 for management abbreviations), and YR represents the duration (years) of management.

2.5 Specific methodology in Chapter 5

Over recent decades farmlands in eastern Africa have been experiencing strong degradation due to the combined effects of agricultural intensification and mismanagement. In Chapter 5 we choose Kenya and Ethiopia as a case study to evaluate the potential impacts of improved agricultural practices on cropland SOC stocks, N loss (gaseous emissions plus N leaching), and crop production. The structure of this chapter includes three parts. First, we test the model performance in simulating the SOC and maize yield response to different managements by comparing with observed data from two long-term field sites in Kenya. Next, country-level yields of six CFTs modelled by LPJ-GUESS are evaluated against FAO-based statistics in Kenya and Ethiopia. In the last part, the isolated effects of each alternative management practice are first investigated for the historical period and subsequently explored under future climate pathways by forcing the model with simulated climate from five general circulation models (GCMs, Eyring et al., 2016). All simulations' spin-up follows the protocol run described in Chapter 3 (see Sect. 2.3 above). The input data used for simulations in this chapter are summarized in Table 2.5, with the specific experimental setups explained in detail below.

Table 2.4 Summary of simulations performed in Chapter 4. See Sect. 2.4 for abbreviations and further explanations.

Purpose of simulation	Code	Trial involved	Time period	Model spin-up	Land-use	Climate	N fertilizer (F) & manure (M)	CC planting (P) & termination (T) date
SOC stocks, N leaching loss, and crop yield evaluation against field-based trials	4-A1	Undersown CCs	$1901\text{-}E_{year}$[a]	500 years; Land use started 470 years PNV, then 30 years of cropland ramp to 1901; CO_2 and N deposition fixed in 1901	$1901\text{-}S_{year\text{-}5}$[a]: Alternating every 5 years between 100% grassland and 100% cropland; $S_{year\text{-}5}$[a] $\text{-}E_{year}$: 100% cropland	GSWP3-W5E5 (daily)	Literature-based	P: 1 d after the previous main crop harvest; T: 1 d before planting the next main crop
	4-A2	CCs grown on fallow cropland	$1901\text{-}E_{year}$[a]	As 4-A1	As 4-A1	GSWP3-W5E5 (daily)	Literature-based	P: 15 d after the previous main crop harvest; T: 15 d before planting the next main crop
Effects of CCs on global crop ecosystems	4-B1	NoCC	1901-2014	As 4-A1	LUH2	CRUJRA (monthly)	F: Ag-GRID; M: Zhang et al. (2017)	As 4-A2
	4-B2	NoCC, CC_L, CC_NL, CC_LNT[b]	2015-2050	Starting from 4-B1 in 2014	LUH2, fixed in 2014	CRUJRA (monthly), 1995-2014 climate repeated until 2050[c]	F: Ag-GRID, fixed in 2014; M: Zhang et al. (2017), fixed in 2014	As 4-A2

a) E_{year}—The year that field experiments ended; $S_{year\text{-}5}$—5 years before the field experiments started; b) NoCC—control treatment with no cover crop; CC_L—legume cover crops; CC_NL—non-legume cover crops; CC_LNT—combined management practice with legume cover crops and no tillage; c) Historical (CRUJRA-based) climate with temperature de-trended. These 20 years are repeated throughout the period 2015-2050.

Table 2.5 Summary of simulations performed in Chapter 5. See Sect. 2.5 for abbreviations and further explanations.

Purpose of simulation	Code	Trial involved	Time period	Model spin-up	Land-use	Climate	Manure input	N fertilizer
SOC storage and maize yield evaluation against field-based trials	5-A1	INM3 site[a]	1901-2015	500 years Land use started 470 years PNV, then 30 years cropland ramp to 1901 CO_2 fixed in 1901	1901-2002: 100% grassland 2003-2015: 100% cropland	GSWP3-W5E5 (daily)	70 kg N ha^{-1}	0, 30, 60, and 90 kg N ha^{-1}
	5-A2	CT1 site[a]	1901-2015	As 5-A1	1901-1991&1995-2000: 100% grassland 1992-1994&2001-2015: 100% cropland	GSWP3-W5E5 (daily)	0 kg N ha^{-1}	0, 30, 60, and 90 kg N ha^{-1}
Regional crop yields comparison & Response to different management practices, historical	5-B1	STD (standard simulation)	1901-2014	As 5-A1, baseline run for 5-B2	LUH2	CRUJRA (monthly)	Zhang et al. (2017)	Ag-GRID
	5-B2	All managements[b]	2015-2100	Starting from 5-B1 in 2014	LUH2, fixed in 2014	CRUJRA (monthly), 1995-2014 climate repeated until 2100 [c]	Zhang et al. (2017), fixed in 2014	Ag-GRID, fixed in 2014
Response to different management practices, future	5-C1	STD (standard simulation)	1901-2014	As 5-A1, baseline run for 5-C2 and 5-C3	LUH2	5 GCMs[d]	Zhang et al. (2017)	Ag-GRID
	5-C2	All managements[b]	2015-2100	Starting from 5-C1 in 2014	LUH2, fixed in 2014	5 GCMs × 3 SSP	Zhang et al. (2017), fixed in 2014	Ag-GRID, fixed in 2014
	5-C3	All managements[b]	2015-2100	Starting from 5-C1 in 2014	LUH2, fixed in 2014	5 GCMs, 1995-2014 climate repeated until 2100[e]	Zhang et al. (2017), fixed in 2014	Ag-GRID, fixed in 2014

a) See Table 2.6 for simulation details; b) All managements denote the seven practices listed in Table 2.7; c) Historical (CRUJRA-based) climate with temperature de-trended. These 20 years are repeated throughout the period 2015-2100; d) Five GCMs—GFDL-ESM4, UKESM1-0-LL, MPI-ESM1-2-HR, IPSL-CM6A-LR, and MRI-ESM2-0—are used, GCM climate is taken from ISIMIP project; e) The same as "c", but GCM-based historical climate.

36

2.5.1 Model evaluation at site scale

To examine model performance, we use data from two long-term experimental sites in western Kenya. The INM3 trial (34.40°E, 0.14°N) mainly evaluates soil fertility effects of manure and maize residue management under tillage systems, while the CT1 trial (34.41°E, 0.13°N) is designed to study the combined effects of conservation tillage and residue application on SOC dynamics in maize systems (Sommer et al., 2018). A total of 16 trials from 2003-2015 in a double-maize cropping system (within a year) at the INM3 site were performed: 0 and 4 t ha⁻¹ of manure dry matter application with 0 and 2 t ha⁻¹ maize residue retention under four levels of N fertilizer addition (0, 30, 60, and 90 kg N ha⁻¹). Similar simulations over the same period were modelled at the CT1 site, but minimum and conventional tillage are dominant practices with no manure application. Double-cropping systems within a year has not been incorporated in LPJ-GUESS since the second "short rainy" growing season, from a yield perspective, is not largely relevant for most parts of eastern Africa. Here, the second growing period in maize-maize systems was simulated as a herbaceous cover crop without N fixation. To parameterize the N application and residue retention in the model, the application rate of 4 t dm ha⁻¹ of manure was set to 70 kg N ha⁻¹ assuming the N content of 1.75% in animal waste (Gichangi et al., 2006). The residue management with 2 t ha⁻¹ application was converted to 50% of maize straw retained in the field, following the proportion reported in Sommer et al. (2018). Furthermore, in the simulations we switched off (on) the tillage option to represent the minimum (conventional) tillage experiment at CT1. A summary of these trials is listed in Table 2.6.

Following the site-level simulation setups in Chapter 3 (Sect. 2.3.1), the gridded climate data set GSWP3-W5E5 was used as input, choosing the grid cell with coordinates 34.25°E and 0.25°N representative for the two trial sites. In terms of land use data in years prior to the field experiments, we followed the simulation setups in Nyawira et al. (2021) and assumed that INM3 was under grassland systems for the period 1901-2002 (5-A1 runs; Table 2.5), while at CT1 grassland was simulated from 1901 to 1991. After this the land use for CT1 trials was implemented given the recorded information in Sommer et al. (2018): rain-fed maize cropping systems from 1992-1994 (unfertilized), followed by a crop-free period of 1995-2000 (grassland), then two years with fertilized maize until 2002 (5-A2 runs; Table 2.5).

To assess the agreement between simulations and observations in SOC stocks, the measured SOC values were scaled from the original depth (0-15 cm) to the modelled soil depth 0-150 cm, using the

depth distribution functions described in Sect. 2.4.3 (also see Eqs. 2.22-2.23). The site-specific empirical parameter—β in Eq. 2.22—was set as 0.971 and 0.974 for INM3 and CT1, respectively, according to the measured values in Nyawira et al. (2021).

Table 2.6 Treatment-specific data used for model evaluation at the INM3 and CT1 trials in Chapter 5.

Site and its soil physical properties	Treatment name[a]	Tillage	Manure (kg N ha⁻¹)	Residue Retention (%)	N fertilizer (kg N ha⁻¹)				
					T[b]	N0	N30	N60	N90
INM3 (34.40°E, 0.14°N)	Nx_NoMan_NoRR	Yes	No	No					
	Nx_NoMan_RR	Yes	No	50					
Topsoil (0-20cm): Sand: 26% Silt: 18% Clay: 56%	Nx_Man_NoRR	Yes	70	No	P	0	10	20	30
Bulk density: 1.1 g cm⁻³	Nx_Man_RR	Yes	70	50					
CT1 (34.41°E, 0.13°N)	Nx_NoTill_NoRR	No	No	No					
	Nx_NoTill_RR	No	No	50					
Topsoil (0-40cm): Sand: 16% Silt: 15%	Nx_Till_NoRR	Yes	No	No	HV	0	20	40	60
Clay: 69% Bulk density: 1.1 g cm⁻³	Nx_Till_RR	Yes	No	50					

a) The "x" in the treatment names denotes any mineral N application rate of 0, 30, 60, and 90 kg N ha⁻¹. Abbreviations: NoMan – no manure application; NoRR – no residue retention; NoTill – no-tillage; Man – 70 kg N ha⁻¹ of manure application converted from 4 t ha⁻¹ dry matter; RR – 50% of residue retention; Till – Tillage; **b)** Abbreviations: T – timing of N fertilization; P – planting date of maize; HV – halfway through the vegetative phase of maize.

2.5.2 Regional crop yields evaluation

In this experiment we only performed one simulation with six CFTs included—maize, pulses, sorghum, wheat, rice, and soybean—which are widely-grown crops in Kenya and Ethiopia (FAOSTAT, 2023). In a previous modelling study (Olin et al., 2015a), sorghum in LPJ-GUESS was simulated as the maize CFT due to the absence of crucial growth parameters (e.g., photosynthetic carbon partitioning) for sorghum. Here, we updated the parameters of assimilate allocation to sorghum organs based on the data from Penning de Vries et al. (1989). The performance of the model for sorghum and five other crops was evaluated by comparing the simulated and reported yields at the national level. For regional comparison, statistics-based yield data were derived from FAOSTAT (2023) while the simulated gridded crop productions were aggregated to country level

using Eq. 2.20 (see Sect. 2.3.2). Regionally, this model simulation (5-B1 run; Table 2.5) was implemented at 0.5° resolution over the period 1901-2014 after model spin-up, with all six CFTs under a conventional management prevalent in eastern Africa (STD; see Table 2.7 for management details). Model forcings—including climate, CO_2 concentrations in the atmosphere, historical land use/land cover, CFTs-specific growth distribution, synthetic N fertilizer and manure application, N deposition, and soil physical properties—were the same as the global simulation setups in Chapter 4 (see Sect. 2.4.2).

2.5.3 Ecosystem responses to management practices in eastern Africa

At the regional level five alternative management practices—legume cover crop (CC_L), non-legume cover crop (CC_{NL}), residue retention (RR), manure application (MAN), and no-tillage (NT)—together with an integrated management were assessed (Table 2.7). The latter, which included most individual practices, was selected to be representing conservation agriculture (CONSERV; see Table 2.7). Simulated outputs of these six management practices were compared with a standard simulation (STD) with setups shown in Table 2.7.

Table 2.7 Simulation setups used for detecting the responses of crop ecosystems to various managements over eastern Africa in Chapter 5.

Simulation*	CC_L	CC_{NL}	RR	MAN	NT	CONSERV	STD
Legume cover crop	Yes[a]	No	No	No	No	Yes[a]	No
Non-legume cover crop	No	Yes[b]	No	No	No	No	No
Residue retention[c]	10%	10%	100%	10%	10%	100%	10%
Manure application[d]	Yes	Yes	Yes	No	Yes	Yes	Yes
Mineral N fertilizer	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tillage	Yes	Yes	Yes	Yes	No	No	Yes

* Abbreviations: CC_L – legume cover crop; CC_{NL} – non-legume cover crop; RR – residue retention; MAN – manure application; NT – no-tillage; CONSERV – conservation agriculture; STD – standard simulation, representing a conventional management prevalent in eastern Africa; **a)** N-fixing C_3 grass; **b)** Competing C_3 and C_4 grasses; **c)** The value of 10% residue retention is taken from the investigated data in farm fields in Ethiopia (Laekemariam et al., 2016; Lemma et al., 2021); **d)** The C:N ratio of farmyard manure is set to 16 in the model simulations, following the literature-reported value in eastern Africa (Gichangi et al., 2006).

The practice that produced the largest SOC increase at each grid cell was chosen as the optimal C management (OPT) for the historical and future simulations:

$$OPT = \{MAX\,(SOC_i - SOC_{STD}), i = 1:5\} \tag{2.25}$$

where OPT is the calculated best-performing C management in a given grid cell; i represents the five management practices of CC_L, CC_{NL}, RR, MAN, and NT. SOC_i and SOC_{STD} are the modelled SOC stocks from these five practices and standard simulation, respectively.

An initial simulation was performed to simulate the effects of these management practices under constant climate, CO_2, and land use in order to isolate effects of management from environmental change. This began with a run of the historical period (1901-2014) after model spin-up, using time-dependent gridded climate, land cover, and N inputs (deposition and fertilizer) at 0.5° resolution. This run generated present-day cropland soil C and N pools under STD management over eastern Africa (5-B1 run; Table 2.5). Subsequent runs, one using each management practice, branched from this present-day state in 2015. In these simulations de-trended climate (repeating 1995-2014) and fixed CO_2 concentration, together with N fertilizer and land cover data of the year 2014, were repeated for 86 years to allow soil C and N pools to reach a new equilibrium after the management shift (5-B2 runs; Table 2.5).

In a second experiment, simulations were driven with future monthly climate data taken from five GCMs, for 1901-2100 at 0.5° spatial resolution (see Table 2.5 for GCMs information). For the historical period (1901-2014), the management setup was the same as the simulation of 5-B1, but with GCM-based climate input (5-C1 runs; Table 2.5). The seven management practices listed in Table 2.7 started in the year 2015, with dynamic climate, CO_2 concentration, and N deposition throughout. Land cover and fertilizer use (mineral N and manure) were fixed from 2014 onwards to exclude their effects on cropland SOC sequestration (5-C2 runs; Table 2.5). N deposition and climate data for SSP1-RCP2.6 (SSP1-26), SSP3-RCP7.0 (SSP3-70) and SSP5-RCP8.5 (SSP5-85) climate change scenarios were selected for future simulations (Meinshausen et al., 2020). Similar to the 5-B2 runs, the long-term management effects excluding environmental change impacts were also investigated using repeated GCM-based climate (5-C3 runs, Table 2.5). The modelled SOC in the last ten years of the 5-C2 simulations (2091-2100) were compared with the 5-C3 outputs over the same period in order to explore the potential transition of the optimal C management (OPT) caused by future climate change. Experimental setup details are available in Table 2.5.

3 Assessment of biological nitrogen fixation in global grain legumes

This chapter evaluates the model performance to simulate BNF-related variables—such as grain yield, legume biomass, soil N uptake, and N fixation rate—in soybean and faba bean using field-based trial data from the literature. Global spatial and temporal patterns of crop yields and BNF rates in these two grain legumes from LPJ-GUESS are subsequently quantified and compared against previously published large-scale estimates.

3.1 Results

3.1.1 Model evaluation at site scale

The overall evaluation of model performance in simulating BNF-related variables across all grain legumes sites is shown in Fig. 3.1. Modelled yields generally agreed well with observations, especially in the site-specific simulation setup. These had higher regression slope (0.83) and lower mean absolute error (28%) compared with the global-uniform simulation setup (Fig. 3.1a). Grain legumes' N content in grains and shoots showed a low agreement, with simulated values underestimating the observations for most sites (Fig. 3.1b-c). This result could be attributed to two important N sources to grain legumes not being captured well by the model (i.e., soil N uptake and BNF, given in Fig. 3.1d-e). The global-uniform run produced a low agreement with observations in N fixation, with a regression slope of 0.22 and mean absolute error of 39%. The simulated underestimation in BNF compared with observations was largely eliminated by using site-specific parameters, with the regression slope increasing to 0.41 and the mean absolute error reducing to 31% (Fig. 3.1e). The field-based measurements showed that the N derived from the atmosphere (%Ndfa) was the main contributor to the legumes' total N uptake, ranging from 15 to 95%, with a mean of 64% across all field trials. LPJ-GUESS in general captured the mean response well, with simulated %Ndfa being 60% and 58% in the site-specific and global-uniform runs, respectively, despite several extreme disagreements at several faba bean sites (Fig. 3.1f).

A linear relationship between legume yields and the rate of BNF was found across a range of field sites in this study (Fig. 3.2a). Simulations from LPJ-GUESS mostly captured the close correlation between these variables, with R^2 ranging from 0.46-0.63 ($p<0.001$) in both runs, not far from the measured value 0.67 (Fig. 3.2a). Also fitted linear regression parameters (i.e., slope and intercept) in both simulations were close to the observations, indicating that the model reproduces well the N fixation effect on yield for individual sites.
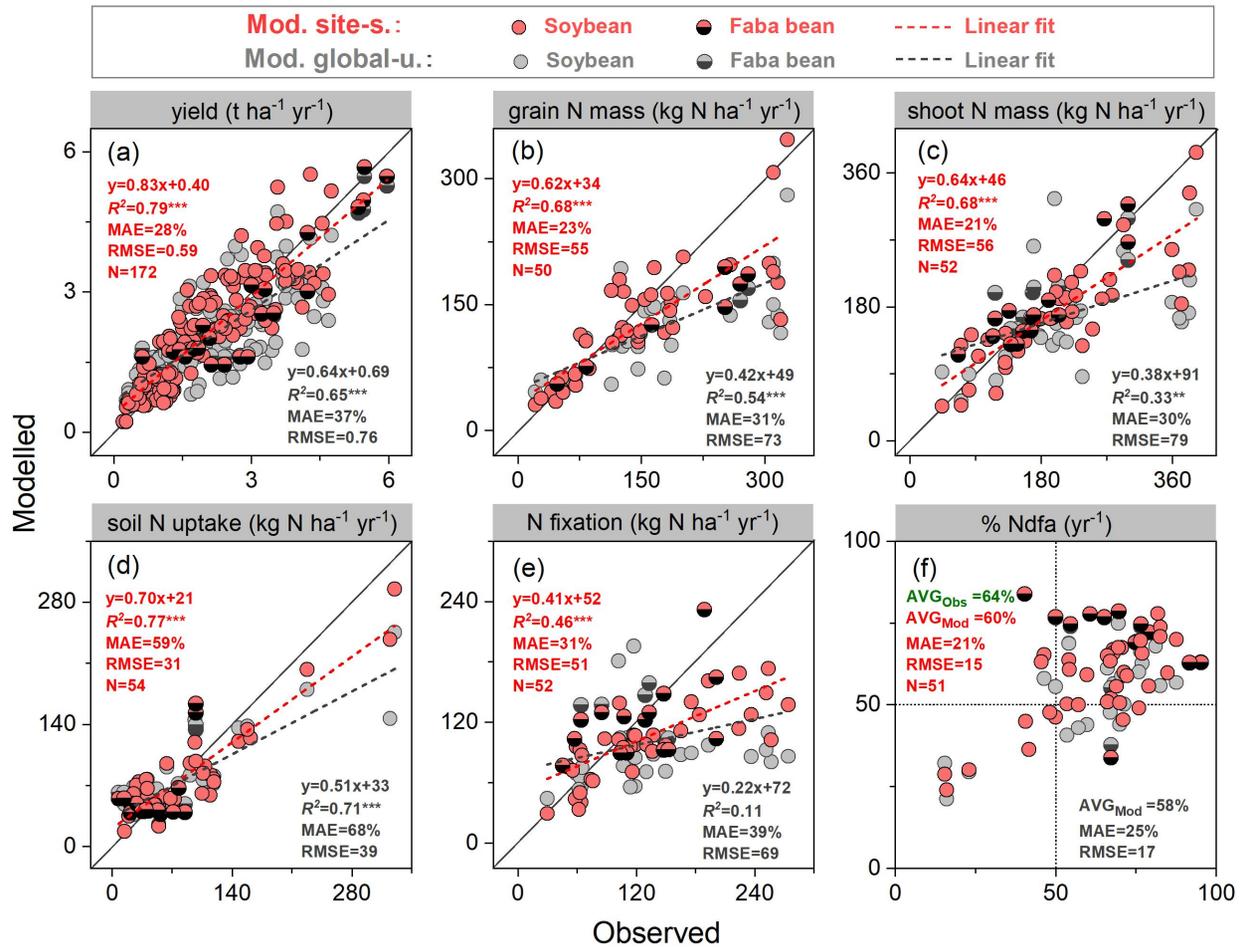
Figure 3.1 Comparison of modelled and observed yield (a), grain N mass (b), shoot N mass (c), soil N uptake (d), BNF (e) and %Ndfa (the proportion of plant N derived from the atmosphere) (f) at harvest across all soybean and faba bean sites. Filled red and grey circles depict 'site-specific' and 'global-uniform' run, respectively. The dashed line is fitted linear regression; *** and ** denote regressions statistically significant at p=0.001 and 0.01, respectively; MAE is mean absolute error, represented in percent (%); the unit of RMSE is the same as the associated variable; AVG in (f) is the averaged value of %Ndfa across all field trials.

A negative exponential relationship was observed between N fertilizer application rate and N fixation when all field trials were included (Fig. 3.2b). LPJ-GUESS reasonably reproduced the decreased trend of BNF to N fertilizer increase, with the similar fitting functions to observations, although higher N fixation rates were modelled in the highest-fertilized trial (600 kg N ha⁻¹) compared with measurements. Additionally, the amount of crop N uptake from soils, as expected, was sensitive to soil mineral N concentration: a significant linear relationship between soil N uptake and N fertilizer application rate (p<0.001) was found in both model simulations and field measurements (Fig. 3.2c).
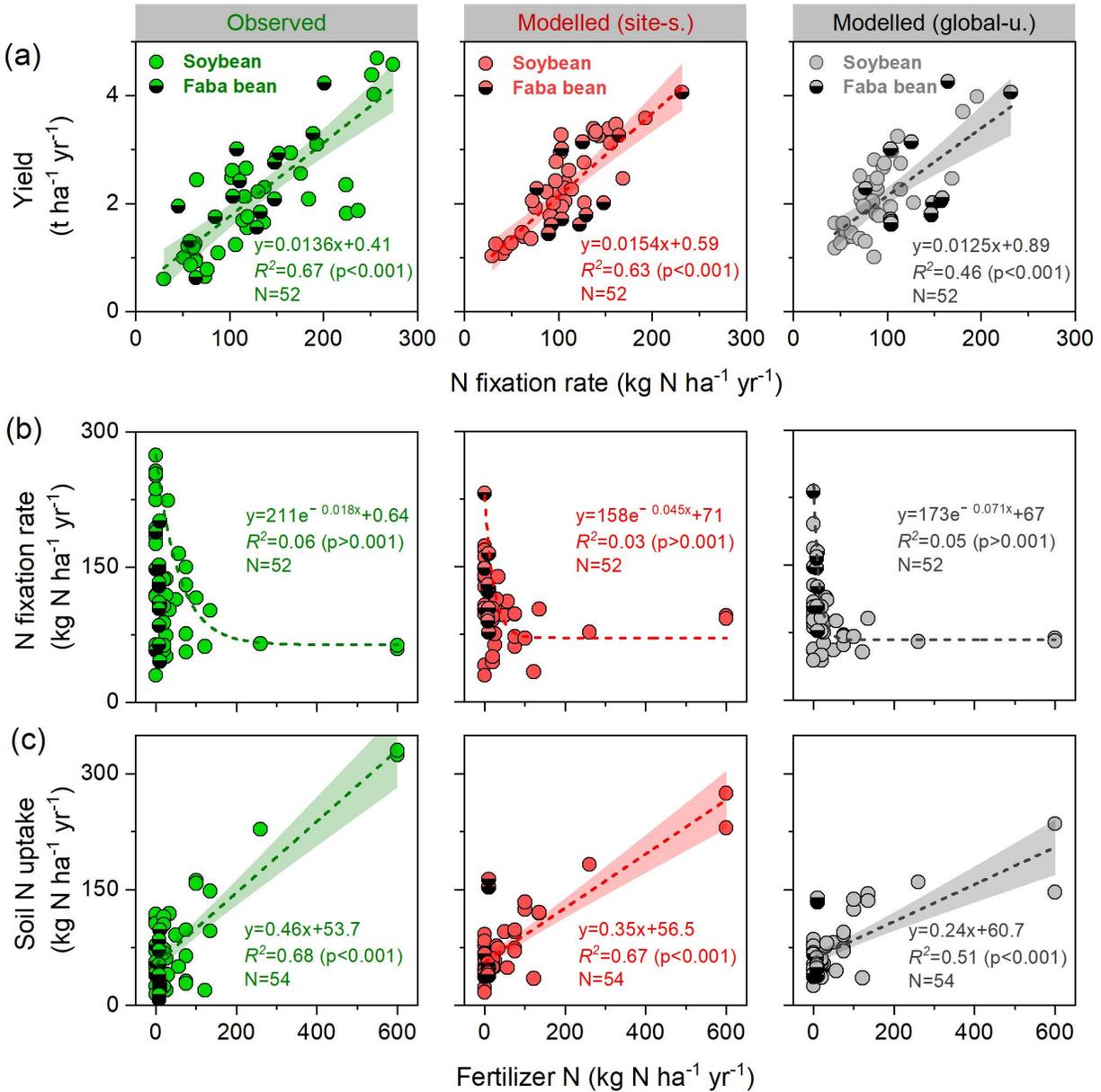
Figure 3.2 Comparison of modelled and observed yield and N fixation rate (a) and the response of BNF (b) and soil N uptake (c) to N-fertilizer addition across all field trials. The shaded areas represent the 95% confidence interval in linear regression.

3.1.2 Global attained yields

Applying the global-uniform parameters introduced in Sect. 2.3.1 (i.e., specific leaf area of soybean and faba bean is set to 40 and 45 m^2 kg^{-1} C, while target grain C:N ratio is represented as a constant of 8 and 10, respectively), together with the time-dependent gridded N fertilizer and manure application data, we modelled soybean and all pulses (adopting the faba bean parameterization, see

Sect. 2.3.2) at a global scale. We mainly calculated data for the period 1996-2005, as the map of crop-specific growth distribution from the MIRCA data set was available for the year 2000.
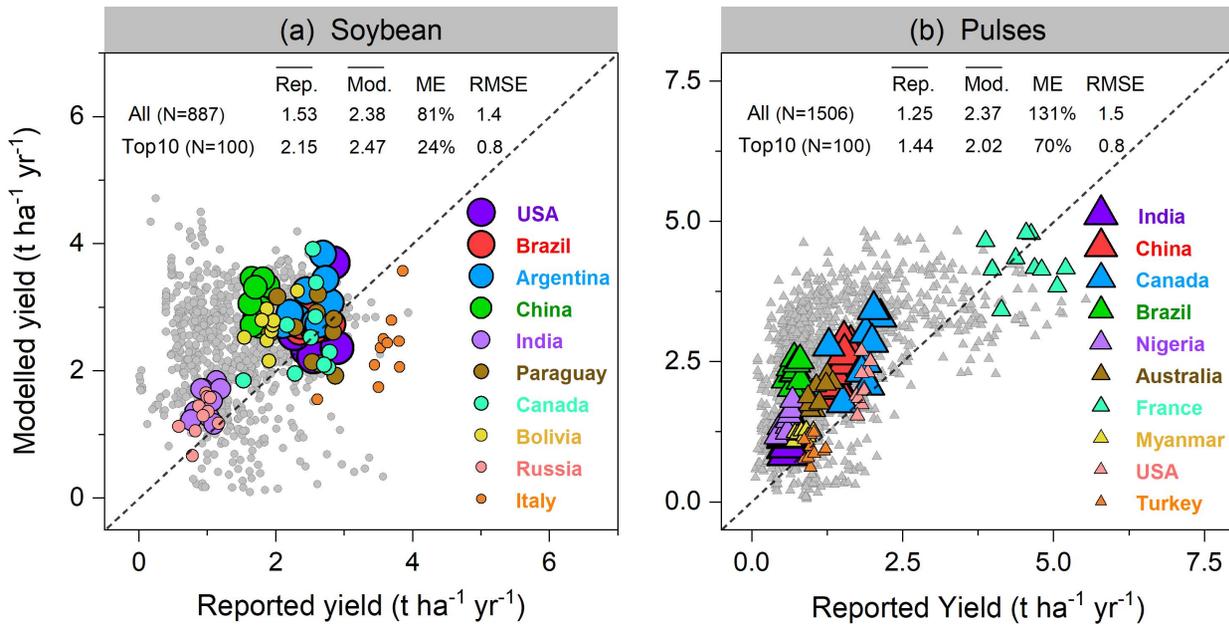
Figure 3.3 Per-country-year comparison of modelled yields of soybean (a) and pulses (b) against FAO statistics from 1996–2005. Each filled circle in (a) represents one year and one country; thus, a country can have up to 10 circles over 1996–2005. In total, 887 and 1506 country-year yield data were used for comparison in soybean and pulses, respectively. The top 10 producer countries shown in color were chosen based on their total production over the same period, and marker size from large to small indicates their total production in descending order. $\overline{\text{Rep.}}$ and $\overline{\text{Mod.}}$ indicate the reported and modelled yields (t ha^{-1} yr^{-1}) averaged from 1996–2005, respectively. ME is mean error, represented in percent (%). RMSE is root mean square error, with the unit of t ha^{-1} yr^{-1} for yield.

Modelled yields in the top ten soybean-producing countries showed a good agreement, with a higher R^2 of 0.52 (p<0.001) and lower RMSE value of 0.8 t ha^{-1} yr^{-1} when low-productivity countries (defined as all countries not belonging to the top ten producer countries) were excluded. With all producer countries included, R^2 of 0.17 (p<0.001) and RMSE of 1.4 t ha^{-1} yr^{-1} was reached (Fig. 3.3). LPJ-GUESS in general tended to overestimate the reported yield for most countries where soybean production is low (e.g., most African countries), with a mean error in such countries ranging from 100-300% (Fig. 3.4). Modelled low yields were found in some arid and semi-arid countries (e.g., Egypt, Iran, Turkey, and Syria), with a range of underestimation from 10-70% (Fig. 3.4). Similarly, overestimated yields were found when comparing simulated yields using the faba bean parameterization against FAO-based report values for pulses in general, with an overestimation also visible for some of the top producing countries (Fig. 3.3). The higher yields

modelled by LPJ-GUESS are most likely resulting from the fairly high N fixation capacity simulated with the faba bean parameterization (see Sect. 3.1.3 below), as well as the widespread distribution of pulses worldwide, which grow under a broad range of climate and soil conditions.

A good fit of the interannual variability of modelled and reported yields is a further indicator of model performance. Despite the deviation between simulations and observations for individual years, modelled variation in soybean production over the period 1981-2016 matched well with FAO-based statistics among the top ten producer countries—particularly in Argentina, India, and China—with a fair Pearson correlation coefficient (r) around 0.60 (p<0.001) and similar standard deviations (Fig. 3.4). The degree of yield variability between years was larger than seen in the FAO records, especially in the U.S., Canada, and Italy (Fig. 3.4), indicating high sensitivity of simulated soybean production to changing environmental factors on spatial patterns, such as weather, N fertilizer application rates, and climate-related N fixation.

3.1.3 Global N fixation and %Ndfa

The spatial pattern of soybean N fixation modelled by LPJ-GUESS presented large discrepancies (Fig. 3.5). Simulated BNF rates as high as 250 kg N ha^{-1} yr^{-1} were found in western South America and most parts of Africa, where neither water nor temperature were critical limitations for N fixation. Moreover, the relatively low fertilizer application in Africa leaves a nitrogen deficit that leads to increased N fixation rates. Conversely, in arid and semi-arid regions, soil water constrains BNF, while temperature limitation is seen in high latitudes and alpine areas (e.g., Andes in Peru). BNF rates in most regions (South Asia, West Asia, Sub-Saharan Africa and northwest China) were as low as 50 kg N ha^{-1} yr^{-1}, particularly in Pakistan and northern India, where simulated BNF is severely constrained by the extreme high temperature over the cropping season (Fig. 3.6). Eastern United States, Europe, Southern China and central-west Brazil showed intermediate fixation rates, which were greater than 150 kg N ha^{-1} yr^{-1}. In general, the spatial variation of modelled legume BNF rate reflects to large degree the spatial climate patterns, in addition to N fertilizer investment. The low simulated %Ndfa of 45±3% in East Asia might indicate high soil N uptake by crops in response to substantial fertilizer application in China over the past four decades. In contrast, the modelled %Ndfa in Africa—with lower N fertilizer use—was as high as 70±3%, despite still lower than the reported mean value of 77% (Table 3.1). The spatial response of N fixation rate to climate constraining factors (i.e., soil temperature and water) is shown for pulses in Fig. 3.6.

We partitioned the simulated grid cells into ten continents described in Sect. 2.3.2 to better compare modelled results with the findings from Peoples et al. (2009). Overall, at regional scale, the simulated outputs showed a good agreement with N fixation rates from the literature (Fig. 3.5). For example, in South America and North America, both major soybean production regions, simulated BNF rates were 156±14 and 127±44 kg N ha^{-1} yr^{-1} over the period 1981-2016, respectively, compared with literature-derived values of 136 and 144 kg N ha^{-1} yr^{-1} (Peoples et al., 2009). Globally, the modelled soybean N fixation rate of 132±21 kg N ha^{-1} yr^{-1} was close to the meta-analysis result of 111-125 kg N ha^{-1} yr^{-1} from Salvagiotti et al. (2008), but lower than a recent estimate of 200 kg N ha^{-1} yr^{-1} based on country-level FAO statistics (Herridge et al., 2022). The contribution of N fixation to total N uptake in soybean was somewhat underestimated in several regions. A similar trend to underestimate reported %Ndfa was also found for pulses (Table 3.1).

Due to large soybean planting areas and high productions, South America and North America were two leading continents fixing N from the atmosphere and jointly contributed 80% of simulated total N fixation in soybean, followed by East Asia, South Asia and Europe (Table 3.1). Globally, the amount of annual N fixed by soybean in model simulation was 11.6±2.2 Tg N over the period 1981-2016, which showed a fairly good agreement with the estimate of 16.4 Tg N yr^{-1} reported by Herridge et al. (2008) and the extrapolated result of 10.4 Tg N yr^{-1} provided by Gelfand & Robertson (2015) based on the U.S. field trials. In addition, our simulated total pulses can fix 5.6±1.0 Tg N from the atmosphere every year, which is about two times higher than an earlier estimate of 2.95 Tg N yr^{-1} (Herridge et al., 2008), but close to a recent finding of 7.2 Tg N yr^{-1} (Herridge et al., 2022). The difference between the estimates in Herridge et al. (2008) and our study is likely due to the low BNF rate used for calculation in Herridge et al. (2008), ranging from 23-86 kg N ha^{-1} yr^{-1}, far lower than the mean value of 119±15 kg N ha^{-1} yr^{-1} in model simulations (Table 3.1).
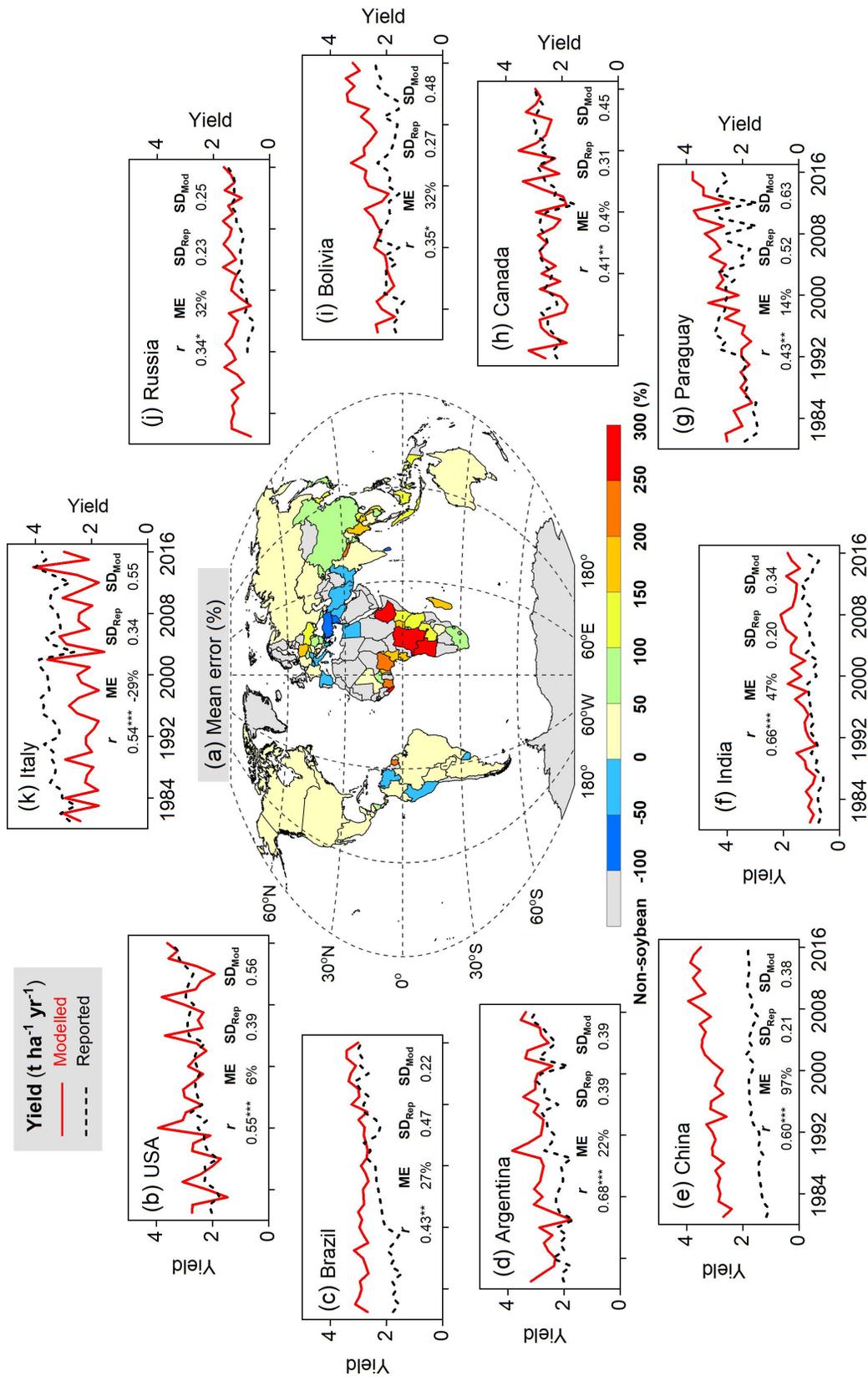
Figure 3.4 Comparison of simulated and FAO-reported yields on country level averaged over 1996–2005 (a), as well as time series of modelled soybean yield (red solid line) and FAO statistics (black dashed line) in the top 10 producer countries (b–k, in descending order) were chosen based on their total production from 1996–2005. r is Pearson correlation coefficient, where ***, ** and * denote the correlation to be statistically significant at p=0.001, 0.01, and 0.05 level, respectively. ME is mean error, represented in percent (%). SD_{Rep} and SD_{Mod} indicate reported and modelled yield standard deviations (t ha^{-1} yr^{-1}) from 1981–2016, respectively.
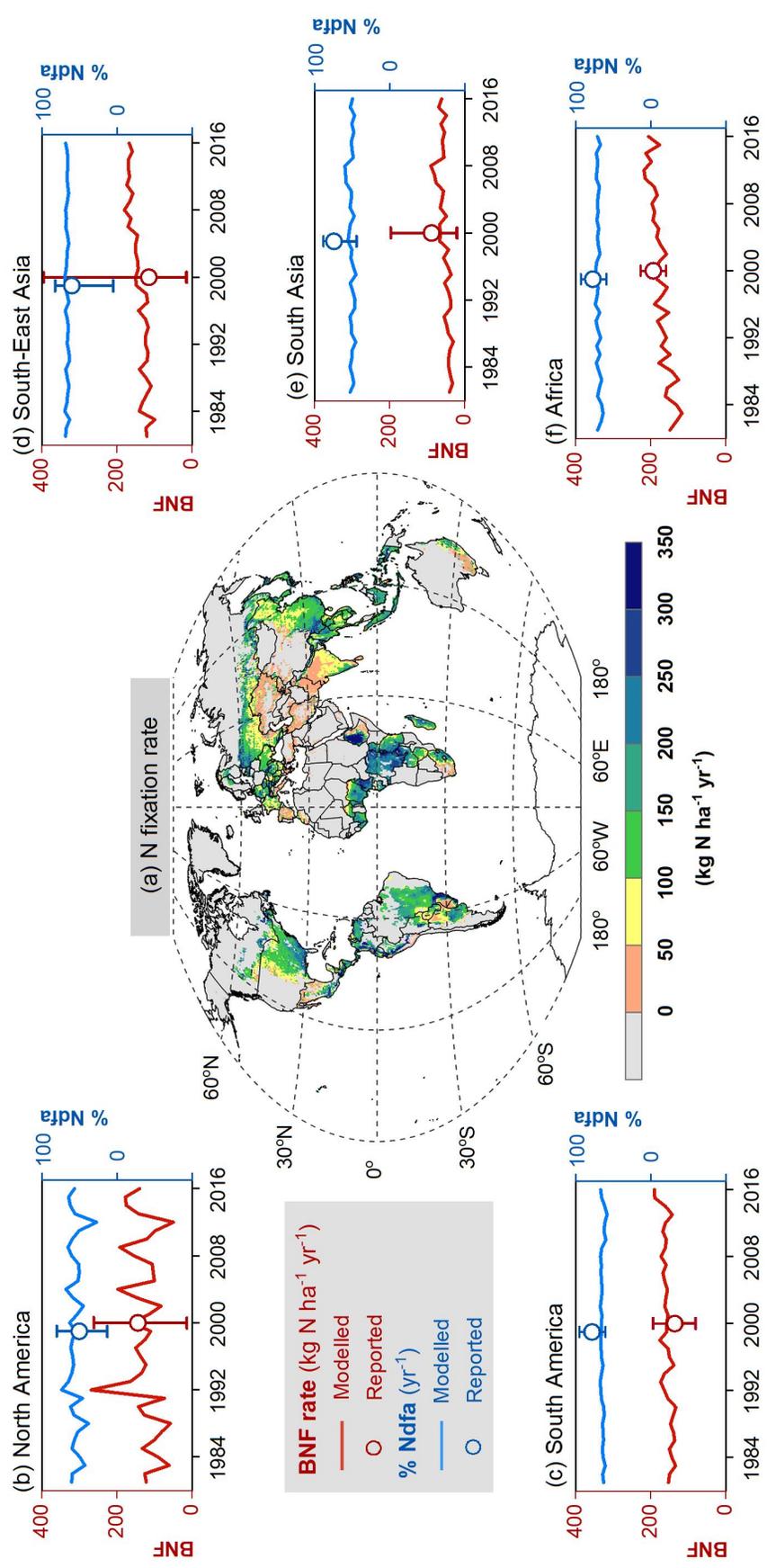
Figure 3.5 Map of soybean N fixation rate modelled by LPJ-GUESS averaged over 1996–2005 (a), and the comparison of simulated BNF rate (red line) and %Ndfa (blue line) with literature-reviewed data (open circle; Peoples et al., 2009) on regional level (b–f). Reported data shown in open circles do not represent any specific years but the potential over time in Peoples et al. (2009), the vertical bars denote the range of estimations based on original literature summarized by Peoples et al. (2009).
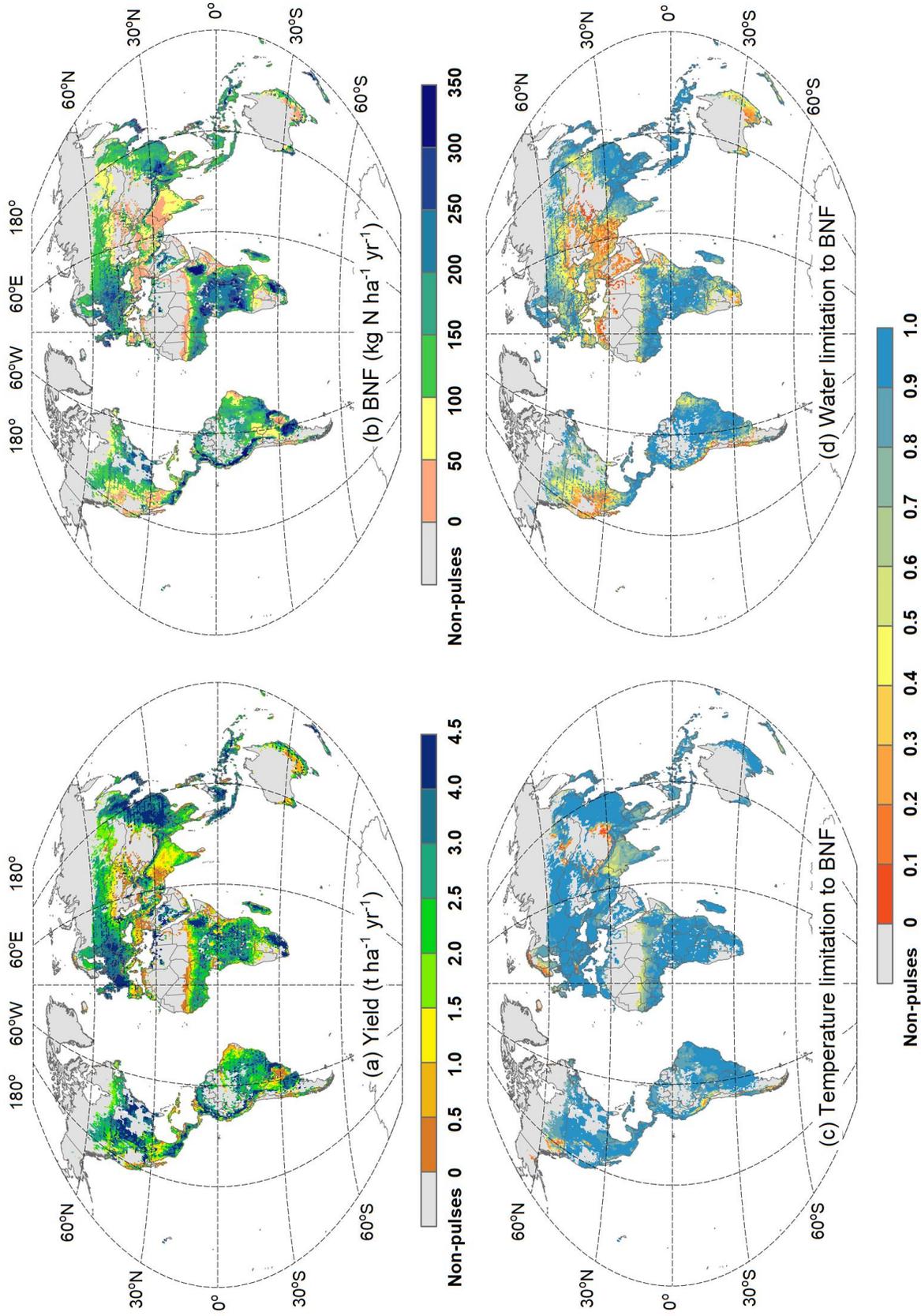
Figure 3.6 Maps of pulses' yield (a) and N fixation rate (b) modelled by LPJ-GUESS, averaged over 1996-2005. The mean soil temperature (f_T; Eq. 2.10) and water limitation (f_W; Eq. 2.11) to N fixation throughout the growing season are presented in (c) and (d), respectively.

49

Table 3.1 Modelled continent-level BNF rate, the proportion of plant N derived from the atmosphere (%Ndfa), and total N fixation in soybean and pulses averaged over 1981–2016, compared to estimates from the literature. The modelled results are represented as mean ± 1 standard deviation.

	Soybean					Pulses				
	BNF rate ($kg\ N\ ha^{-1}\ yr^{-1}$)		%Ndfa (yr^{-1})		Total N fixation ($Tg\ N\ yr^{-1}$)	BNF rate ($kg\ N\ ha^{-1}\ yr^{-1}$)		%Ndfa (yr^{-1})		Total N fixation ($Tg\ N\ yr^{-1}$)
	Reported	Modelled	Reported	Modelled	Modelled	Reported	Modelled	Reported	Modelled	Modelled
South Asia	88[a] (21-197)	53±14	74[a] (44-88)	51±3	0.4±0.1	–	62±12	–	52±2	0.9±0.2
South-East Asia	115[a] (0-400)	141±22	60[a] (0-82)	66±2	0.2±0.0	–	139±16	–	69±1	0.3±0.1
Africa	193[a] (159-227)	172±25	77[a] (65-89)	70±3	0.2±0.1	–	157±21	-	70±1	1.9±0.5
North America	144[a] (14-311)	127±44	50[a] (13-80)	56±9	4.9±1.7	118[f] (13-252)	137±21	74[f] (60-92)	59±3	0.6±0.1
South America	136[a] (80-193)	156±14	78[a] (60-95)	64±2	4.5±1.1	–	157±18	–	66±3	0.5±0.1
East Asia	–	101±16	–	45±3	1.2±0.2	–	114±17	–	49±4	0.4±0.1
Central Asia	–	63±19	–	36±6	0.0±0.0[e]	–	104±21	–	59±4	0.0±0.0[e]
West Asia	–	27±7	–	14±3	0.0±0.0[e]	100[f] (78-133)	65±10	69[f] (63-76)	35±4	0.2±0.0
Europe	–	117±17	–	54±4	0.2±0.1	153[f] (73-211)	177±26	74[f] (60-92)	63±4	0.6±0.1
Oceania	–	78±27	–	38±9	0.0±0.0[e]	143[f] (82-216)	126±23	82[f] (69-89)	37±6	0.4±0.1
Global	111-200[a,b,c,d]	132±21	52-68[a,b,c,d]	57±4	11.6±2.2	107-143[c,f,g]	119±15	71-75[c,f,g]	60±1	5.6±1.0

a) Soybean data in Peoples et al. (2009); **b)** Salvagiotti et al. (2008); **c)** Herridge et al. (2008); **d)** Soybean data in Herridge et al. (2022); **e)** The values do not represent zero grain legume planting area in that region; **f)** Faba bean data in Peoples et al. (2009); **g)** Faba bean data in Herridge et al. (2022).

3.2 Discussion

3.2.1 Model performance at site scale

The simulated legume yields and grain N mass at harvest were on average 20-30% lower than values reported in the measurements across a range of field sites (Fig. 3.1). A similar, small, underestimation was found in the shoot N mass (Fig. 3.1), indicating that the productivity generally is somewhat too low in the model. One factor contributing to the underestimation is that LPJ-GUESS applies a conversion factor of 2.0 from plant C mass to dry matter, ~10% lower than the published measurement of 2.24 reported in Osaki (1993). Additionally, we found that the model underestimated aboveground biomass while simultaneously overestimating belowground productivity at the three sites where measured root biomass was available (not shown). This could be addressed by adjusting the root:shoot allocation (Eq. 2.5), but this is currently prevented by the insufficient observed root biomass information.

Modelled soil N uptake was sensitive to soil mineral N concentration and hence driven by fertilizer application rates (Fig. 3.2). Generally, LPJ-GUESS tended to underestimate soil N uptake in regions where legumes were well-fertilized (Fig. 3.2). This might be partially due to the saturation effect of mineral N concentration on N uptake implemented in the model, which can cause the discontinuation of N uptake when soil available N is abundant (Zaehle & Friend, 2010; Wårlind et al., 2014). Under high fertilization rates, a strong underestimation in soil N uptake was expected due to the simulated saturation-response to high soil mineral N, resulting in little changes in the level of soil N uptake no matter how much N fertilizer was applied.

The percentage of plant N derived from the atmosphere (i.e., %Ndfa) is a key parameter required for quantifying N fixation in the field (Peoples et al., 2021). LPJ-GUESS captured the range and mean value of %Ndfa well across different field trials despite some disagreements between simulations and observations for individual experiments (Fig. 3.1). An underestimated %Ndfa is likely a result from the combined effects of underestimated N fixation and overestimated soil N uptake. Nevertheless, we found modelled %Ndfa to decline with increasing N fertilizer application, which is consistent with the observed response in the field trials. A negative correlation between %Ndfa and fertilizer application rates was also reported by Salvagiotti et al. (2008) based on a range of reviews and experimental papers. These results all suggest that LPJ-GUESS is able to effectively capture the observed overall patterns of soil mineral N uptake and N fixation in grain legumes and their responses.

Since the SLA and C:N ratio of plant organs are of importance in determining N uptake when modelling vegetation C-N dynamics (Camargo-Alvarez et al., 2022), it is expected that applying measured values for site-scale simulation resulted in much better agreement when comparing modelled results with measurements (Fig. 3.1). Remaining disagreement between simulated and observed N variables may reflect missing processes in the model, such as inoculation effectiveness and phosphorus limitation, particularly in terms of inoculant application. Field experiments indicate that proper inoculation of rhizobia promotes nodulation and results in an efficient increase in N fixation (Mínguez et al., 1993; Tewari et al., 2004; Denton et al., 2017). Using a fixed parameter ($N_{maxfixpot}$; Eq. 2.9) to represent all inoculation situations such as in global-uniform calibration cannot reflect this variability. Moreover, due to the difficulties in measuring both nodules and roots under field condition, in many studies the observed BNF rates were determined by legumes' aboveground biomass. Excluding the root contribution to the whole plant BNF rates likely cause an underestimation of N fixation: N associated with nodules and roots in soybean and faba bean may account for 20-40% of the total N accumulation at mid-flowering phase (Unkovich & Pate, 2000; Khan et al., 2003).

3.2.2 Global yields, N fixation, and %Ndfa

In some arid and semi-arid countries, the simulated national-level soybean yields were up to 70% lower than FAO-based statistics likely in response to the simulated low BNF rate caused by water constraints (Fig. 3.6). Conversely, simulations from LPJ-GUESS overestimated crop production by 100-300% in some African countries, with modelled BNF rates of 300-350 kg N ha^{-1} yr^{-1} (Fig. 3.5). More recent studies from African farms have suggested that the soybean N fixation rate can be as low as 0-50 kg N ha^{-1} yr^{-1} in most farmer fields, due to the inconsistent effectiveness of inoculation in the acid soils (Ulzen et al., 2016; Vanlauwe et al., 2019). Unfortunately, current BNF implementation and soil representation in LPJ-GUESS do not account for inoculation effectiveness in response to soil pH.

In our simulations, the annual amount of N fixed by soybean and all pulses of 17.2±2.9 Tg averaged over the period 1981-2016 agreed well with the estimate of 19.4 Tg provided by Herridge et al. (2008), in which the crop production statistics from FAO and legume-specific %Ndfa from farmer fields were used for estimating global N fixation. In an earlier study, a total of 10 Tg N was estimated from legume crop BNF annually (Smil, 1999), far lower than our findings. The discrepancy between the estimates in Smil (1999) and Herridge et al. (2008) likely reflect the lower

values of % Ndfa for soybean and pulses used for calculations in Smil (1999). Also, Smil (1999) excluded belowground fixed N associated with roots and nodules, which contributes to the low estimate. Our modelled N fixation from grain legumes amounts to 12% of the annual mean of ca. 140 Tg N that were estimated to be fixed across global terrestrial ecosystems (Cleveland et al., 2013;Vitousek et al., 2013; Meyerholt et al., 2016; Davies-Barnard & Friedlingstein, 2020; Davies-Barnard et al., 2022), indicating the importance of BNF input in croplands for the global terrestrial N cycle, although large percent of the fixed N is removed by grain yields from the ecosystems annually.

At present three environmental factors, soil temperature, moisture, and soil mineral N concentration, affect N fixation in the model. As discussed above, increased soil N concentration would depress N fixation as plant N demand can be fulfilled more 'cheaply' via soil mineral N uptake. This effect is also visible from the spatial pattern of %Ndfa in the northern temperate region, such as the United States, western Europe, and China. Here, anthropogenic N deposition, together with the intensive application of fertilizers result in soils being N-rich, inhibiting simulated BNF. This could explain why our modelled soybean N fixation rate was not high in East Asia. In comparison, the high rate of N fixation found in tropical regions is mostly due to their high nitrogenase activity under warm and moist soil conditions, resulting in %Ndfa of ~70% being modelled for all grain legumes in the tropics (Table 3.1). A similar spatial variation between temperate and tropical regions in N fixation was also reported by other modelling studies (e.g., Wang & Houlton, 2009; Meyerholt et al., 2016; Xu-Ri & Prentice, 2017). Taken together, these results reveal that LPJ-GUESS broadly captures how N managements and climate variation affect soil N uptake and N fixation in grain legumes at large spatial scales.

Although N fixation can help grain legumes to significantly enhance their total N accumulation and to achieve higher N concentration in seeds, these benefits are usually accompanied by the increase in respiration cost of 8-32% of net photosynthetic carbon (Kaschuk et al., 2009; 2010). Such a respiratory photosynthate consumption would reduce productivity if photosynthesis rate does not increase to compensate for the cost. In LPJ-GUESS, we assumed that up to 50% of daily NPP can be used for N fixation. This approach has the advantage that legumes are able to seek to maximize photosynthetic gain due to reduced N limitation in leaf carboxylation capacity, but it entails the risk of lower productivity if too much NPP is invested into fixation. Nevertheless, we did not observe any modelled extreme C expense on N fixation over the entire growing season, where soybeans

were usually spending 5-25% of daily NPP on fixation worldwide (Fig. 3.7). Such NPP consumption was not only lower than our assumed upper limit of 50%, but also appropriately in line with the reported range of 8-32% in Kaschuk et al. (2009), demonstrating the reasonable C cost scheme implemented for N fixation in our model. Taken together, the modelled C profits due to N fixation can be attributed to the positive feedback between BNF and photosynthesis in LPJ-GUESS: C cost-based N fixation results in a higher rate of photosynthesis due to the enhanced leaf N content; in turn, the increased rate compensates for the C cost, and allocates more assimilate to roots and thus enhances N fixation.
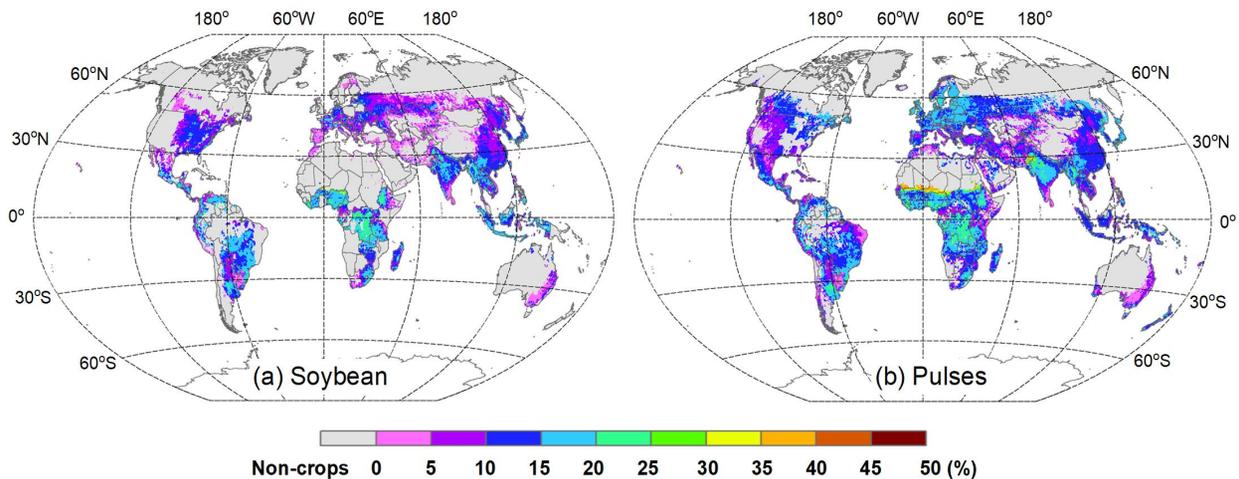
Figure 3.7 Map of daily NPP used for N fixation (%) in soybean (a) and pulses (b) modelled by LPJ-GUESS, averaged over 1996-2005 throughout the growing season. Here we assumed that at maximum 50% of daily NPP can be used for N fixation (Eq. 2.15).

3.3 Conclusions

In this chapter we implemented a process of symbiotic biological N fixation in grain legumes into the crop module of LPJ-GUESS. The modelled C-N variables of soybean and faba bean were extensively evaluated with observed data from site scale to a large region. Our results showed that the BNF scheme adopted in LPJ-GUESS realistically responded to water and N managements, as well as to climate variation, and produced N fixation and yields which generally agreed with measurements.

Our model estimated that global biological N fixation in soybean and total pulses was 17.2±2.9 Tg N yr^{-1} during the period 1981-2016 and that the highest fixation rate occurred in tropical and temperate regions with warm and moist climate. Soil water and temperature were dominant controls on N fixation, in addition to N fertilizer application. On a global average, fixing N from the

atmosphere was the main source to meet the legume N demand, contributing 57±4% and 60±1% to total N uptake in soybean and pulses, respectively. However, processes missing from the model, such as inoculation effectiveness and soil acidity, might have biased estimates on N fixation and yields at a global scale.

Combining the N dynamic process of N fixation with a C-N allocation scheme for crops in LPJ-GUESS provides an opportunity to estimate the changes in global grain legume production and global terrestrial C and N pools under future land use or climate change scenarios. It can also help to predict and detect the potential contribution of N-fixing plants as "green manure" to benefiting soil fertility and agricultural production in global croplands.

4 Global influence of cover crops on yield and cropland carbon and nitrogen balances

This chapter explores the potential contribution of herbaceous N fixers to the sustainable development of agricultural production based on the updated LPJ-GUESS model version from Chapter 3. The objective of this section is to assess and compare the effects of two cover crop functional types—leguminous (CC_L) and non-leguminous (CC_{NL}) grasses—on SOC stocks, N leaching loss, and agricultural productivity across four cropping systems globally (i.e., maize, wheat, rice, and soybean). These three simulated ecosystem service indicators are extensively examined with worldwide site-level observed data and compared against global-level estimates from the existing literature.

4.1 Results

4.1.1 Model evaluation at site scale

Modelled SOC generally agreed well with observations, with high regression slopes (0.75-0.81) and low absolute errors (13-15%) in the control (i.e., NoCC) and cover crop treatments (Fig. 4.1a). We found enhanced cropland soil carbon stocks in the two simulated cover crop types compared with NoCC, indicated by positive annual SOC sequestration rates of 0.28 and 0.45 Mg C ha^{-1} yr^{-1} (on average) in the CC_{NL} and CC_L simulations, respectively (Table 4.1). These compared well with the observed values of 0.78 and 0.48 Mg C ha^{-1} yr^{-1} although the model underestimated the soil carbon enhancement (the range between the 5th and 95th percentiles) when all cover crop types were included (the ranges of -2.1 to 17.2% and 0.8 to 5.8% for observations and simulations, respectively; Fig. 4.1b).

Simulated N leaching from bare fallow cropland (NoCC) tended to be somewhat lower than the measurements, with a mean underestimation of 14%. In contrast, the model overestimated N losses by 57% in the cover crop experiments (Fig. 4.1a). A positive exponential relationship between N fertilizer rate and N leaching ($p<0.01$) was observed across a range of field sites in this study (Fig. 4.2). Simulations from LPJ-GUESS mostly captured this relationship, although higher leached N rates were modelled in the high fertilized trials (224-260 kg N ha^{-1}) compared with measurements (Figs. 4.1-4.2). Replacing bare fallows with cover crops on average reduced N leaching by 54% in the field experiments, with the decreases ranging from 20-87% for non-legume types and 40-68% for legume types (Table 4.1). LPJ-GUESS reproduced these mean differences well, but

underestimated the relative changes in response to cover crops, with the modelled reduction of 5-53% and 4-65% in the CC_{NL} and CC_L simulations, respectively (Table 4.1; Fig. 4.1b).
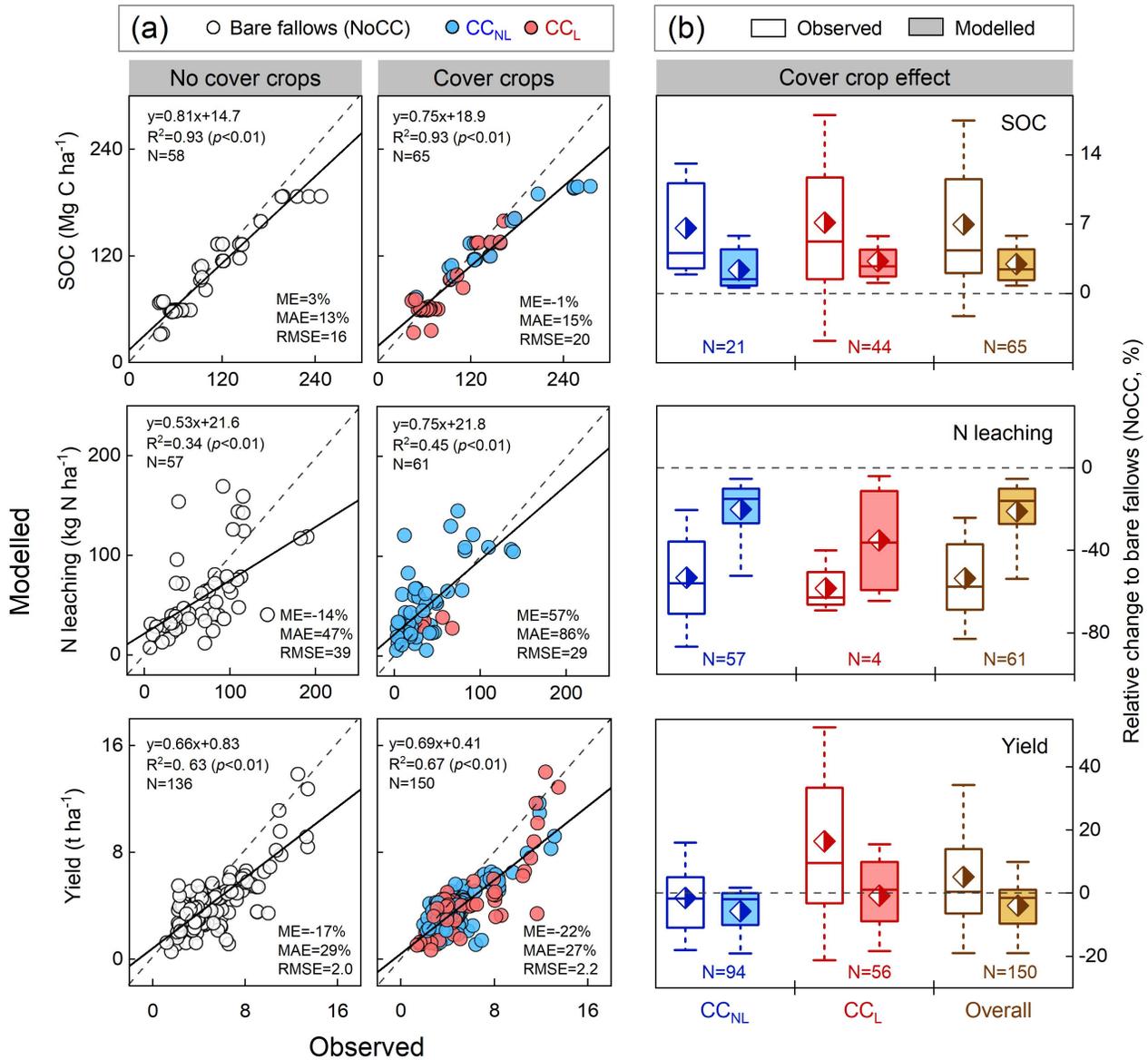
Figure 4.1 Comparison of modelled and observed cropland SOC stocks, N leaching and crop yield (a) and their responses to CCs (b) across all field trials. The dashed line in (a) is the 1:1 line and the black bold line is a fitted linear regression; ME and MAE indicate mean error and mean absolute error, respectively, representing in percent (%); RMSE is root mean square error, with units of Mg C ha^{-1} for SOC, kg N ha^{-1} for N leaching, and t ha^{-1} for yield. Box plots in (b) denote the 5th and 95th percentiles by the whiskers, median and interquartile range are the box lines, and means are symbolized as diamonds.

In comparison with observations, LPJ-GUESS underestimated crop yields on average by 17-22% across all field trials (Fig. 4.1a), mainly as a result of simulated lower agricultural productivity in

the unfertilized systems, particularly in wheat and rice (Fig. 4.2). Compared with the bare fallows, using non-legume CCs during the off-season was modelled to reduce the subsequent main-crop production by 2-16% across four assessed farming systems, larger than the mean observed yield reductions (1-4%) in the field measurements. However, the implementation of N-fixing CCs in our simulations resulted in yield increases in some cases, with the production changes from -18.0 to 16.0% when all crop types were included, falling within the reported range of -21 to 52% (Table 4.1; Fig. 4.1b). In field experiments the yield increase due to legume CCs was largest in unfertilized systems, and the impact of legume cover cropping gradually declined when main crops received high N application rates. The model reasonably reproduced the decreased trend of yield benefits to N fertilizer increases, but generally underestimated these effects in most N fertilization trials (Fig. 4.2).

Table 4.1 Modelled and observed responses of SOC stocks, N leaching, and crop yields to cover crop management compared with bare fallows (NoCC). The changes are represented as the mean values across all field trials, with a range between the 5th and 95th percentiles given in parentheses.

	Unit	CC$_{NL}$		CC$_L$		Overall	
		Observed	Modelled	Observed	Modelled	Observed	Modelled
ΔSOC_{rate}*	Mg C ha^{-1} yr^{-1}	0.78 (0.13, 1.69)	0.28 (0.05, 0.52)	0.48 (-1.21, 2.19)	0.45 (0.16, 0.60)	0.73 (-0.36, 2.19)	0.38 (0.07, 0.60)
SOC change	%	7.0 (2.1, 13.5)	2.4 (0.6, 5.9)	6.3 (-4.6, 17.5)	3.5 (1.2, 5.7)	7.1 (-2.1, 17.2)	3.0 (0.8, 5.8)
N leaching loss	%	-53.2 (-86.5, -20.4)	-20.0 (-52.2, -5.2)	-58.3 (-67.7, -40)	-35.2 (-64.4, -4.0)	-53.6 (-82.8, -24.1)	-21.0 (-53.0, -5.2)
Wheat yield	%	-0.7 (-19.4, 15.9)	-2.8 (-9.7, 1.3)	-1.1 (-34.1, 30.6)	-15.1 (-48.1, 13.4)	-0.8 (-22.0, 22.3)	-6.2 (-31.3, 2.2)
Maize yield	%	-0.6 (-7.0, 7.5)	-15.9 (-24.2, 0.5)	26.7 (-3.6, 80.4)	1.4 (-8.1, 12.7)	11.8 (-6.3, 26.4)	-10.5 (-23.4, 7.6)
Rice yield	%	-4.1 (-27.8, 24.1)	-12.6 (-24.6, -6.6)	31.1 (7.1, 55.3)	3.8 (-10.9, 17.2)	14.6 (-15.4, 44.4)	-3.9 (-15.7, 14.5)
Soybean yield	%	-3.2 (-15.2, 6.3)	-2.1 (-5.9, 0.8)	2.3 (4.2, 7.9)	1.6 (-3.3, 5.0)	-1.5 (-13.2, 7.9)	-0.5 (-4.1, 3.6)
Yield of all crops	%	-1.6 (-18.0, 16.0)	-5.7 (-19.1, 1.0)	16.4 (-21.1, 52.6)	0.8 (-18.3, 15.5)	5.2 (-19.0, 34.3)	-3.8 (-19.1, 10.0)

* ΔSOC_{rate}: annual SOC sequestration rate, see Eq. 2.24 for calculation details.
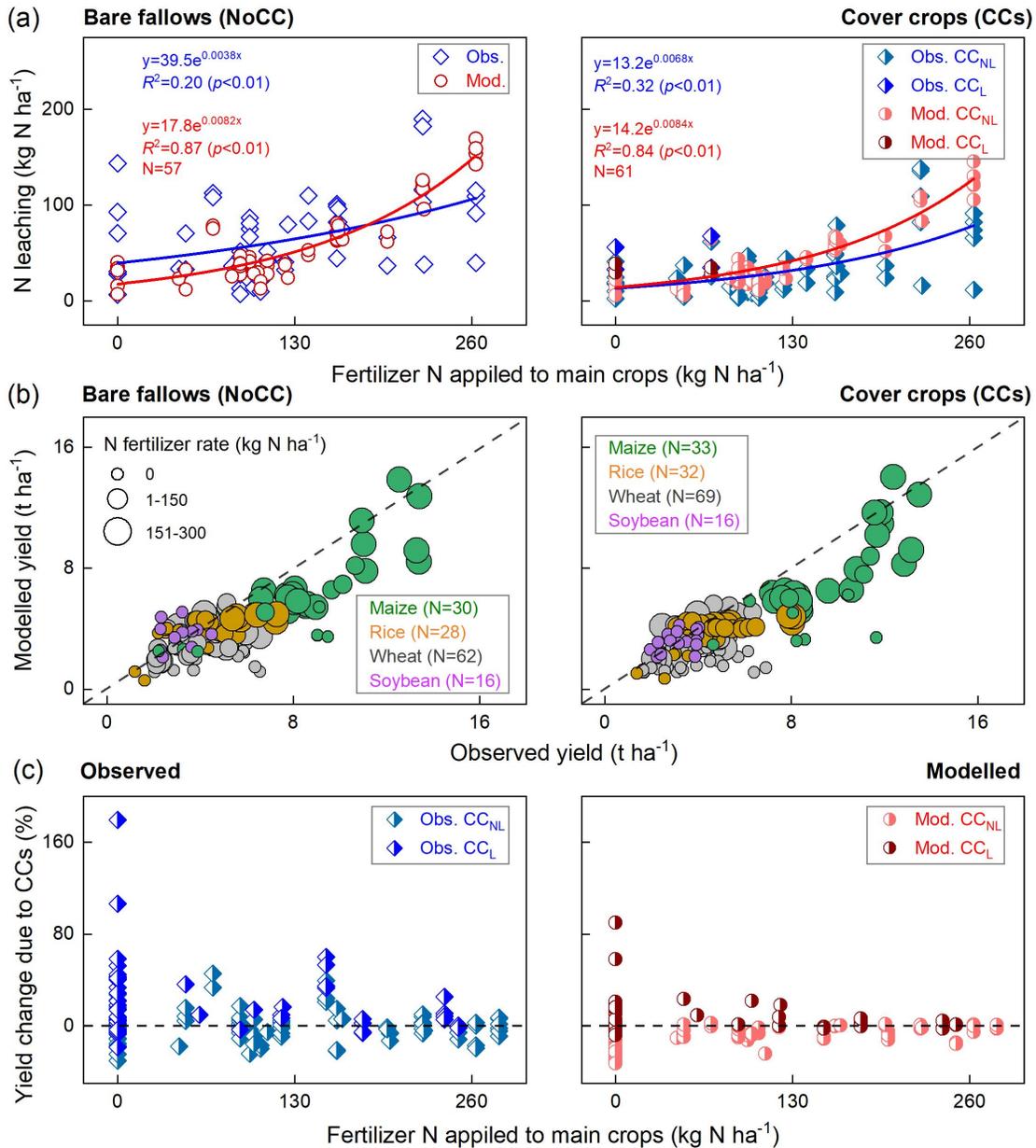
Figure 4.2 Comparison of modelled and observed responses of N leaching (a) and main-crop yield (b) to N fertilizer application across all field trials in the control (bare fallows) and cover crop managements. The modelled and observed influence of cover crops on main-crop production in response to N fertilizer application are shown in (c). Abbreviations: CC_{NL}- non-legume cover crops; CC_L- legume cover crops.

4.1.2 Global crop ecosystem responses to cover crops

4.1.2.1 Soil carbon stocks

Our simulations of the three explored CC managements resulted in a net soil C increase across global croplands compared with the control treatment (NoCC), with the largest SOC sequestration rates (ΔSOC_{rate}; Eq. 2.24) found in warm and moist regions (e.g., Southern China and Eastern

Europe; Figs. 4.3-4.4). For the 36-year simulation period, the maximum annual rates of soil C sequestration in the CC_{NL} and CC_L runs were reached in the sixth year after introducing cover cropping, whereas in the $CC_L NT$ simulation they were already achieved in the fourth year after the implementation of altered management (Fig. 4.4). After these initial peaks, the annual soil C accumulation effect persisted over the course of the remaining simulation period, but with declining rates. On average, using CCs was modelled to sequester 0.10, 0.14, and 0.32 Mg C ha^{-1} yr^{-1} of soil carbon in the CC_{NL}, CC_L, and $CC_L NT$ runs, respectively (Fig. 4.4).
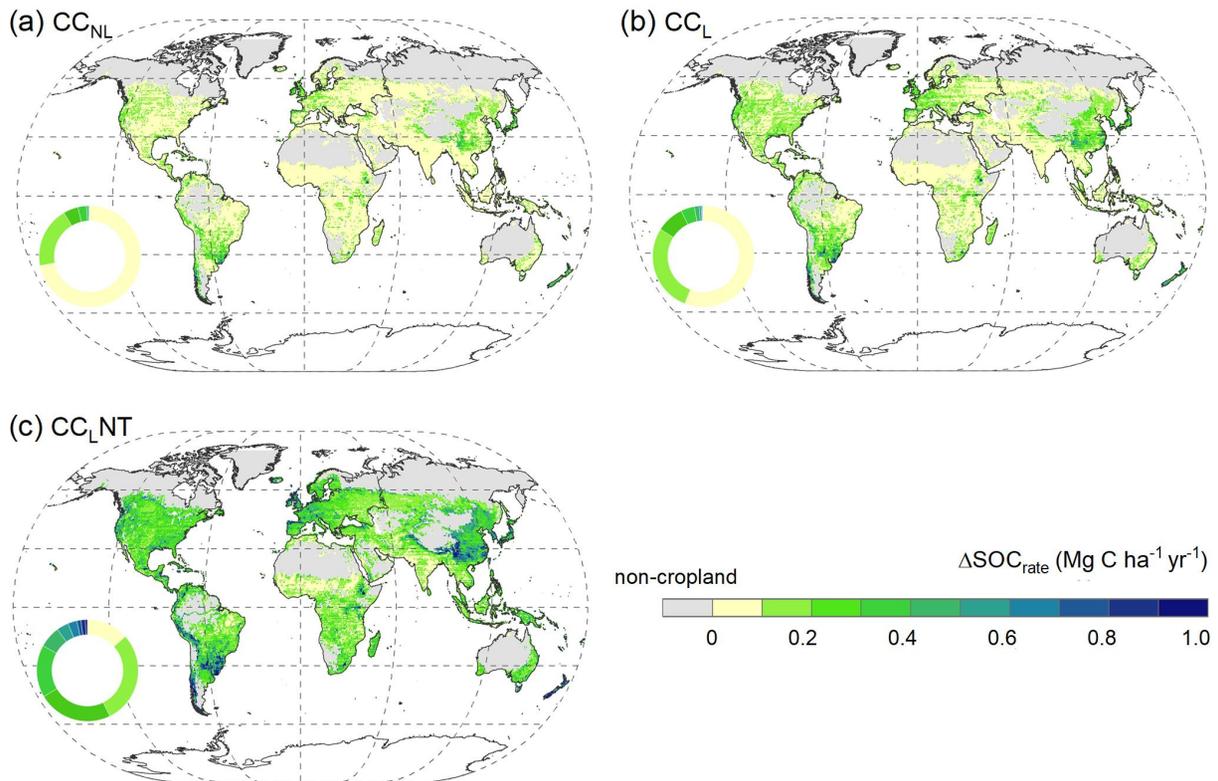
Figure 4.3 Maps of annual soil C sequestration rate (ΔSOC_{rate}; Mg C ha^{-1} yr^{-1}) in response to CC_{NL} (a), CC_L (b), and $CC_L NT$ (c) managements, relative to the control treatment with bare fallows (NoCC) in the 36th simulated year. The inset donut plots represent the area proportion of each classified ΔSOC_{rate} from the total cropland area.

Assuming all current cropland under cover cropping practice worldwide (i.e., CA_{all} scenario), modelled total soil C stocks (0-150 cm) of the various managements ranged from 164.9 to 176.4 Pg C across global croplands, somewhat larger than the published estimates for the topsoil layer (140 Pg C in 0-30 cm, Zomer et al., 2017; 115 Pg C in 0-50 cm, Ren et al., 2020) and within the reported values for the depth 0-100 cm ranging between 157 and 164 Pg C (Jobbágy & Jackson, 2000; Global Soil Data Task, 2014) and 210 Pg C (for 0-200 cm; Jobbágy & Jackson, 2000). In comparison with bare fallows (NoCC), simulations from LPJ-GUESS resulted in an increase of soil

C storage by 3.8 (+2.3%) and 5.4 Pg C (+3.3%) after 36 years of implementation of non-legume (CC_{NL}) and legume cover crops (CC_L), respectively, between the main cropping seasons. Adopting no tillage (CC_LNT) further contributed to increasing simulated soil C storage by 11.5 Pg C (+7.0%) in global croplands (CA_{all} scenario; Table 4.2).

Table 4.2 Modelled total cropland SOC stocks (0-150 cm), N leaching loss, and crop production with alternative cover crop managements under three CA area scenarios in the first and last simulated decades, compared with literature-based estimates. See Fig. 2.4 for spatial pattern of three CA area scenarios.

Scenario	Management	Soil C stock, total (Pg C)		N leaching, total (Tg N yr^{-1})		Crop production[a] (million tonnes per year)	
		1-10 years	27-36 years	1-10 years	27-36 years	1-10 years	27-36 years
CA_{his} (126 Mha)	NoCC	15.8	15.6	0.88	0.80	301	287
	CC_{NL}	15.9	15.9	0.52	0.49	286	292
	CC_L	16.0	16.1	0.58	0.54	295	306
	CC_LNT	16.2	16.7	0.51	0.48	279	294
CA_{pot} (590 Mha)	NoCC	68.9	68.0	5.2	5.4	1145	1126
	CC_{NL}	69.4	69.5	3.3	3.2	1068	1119
	CC_L	69.6	70.2	3.7	3.6	1105	1172
	CC_LNT	70.3	72.5	3.2	3.2	1034	1125
CA_{all} (1597 Mha)	NoCC	167.3	164.9	18.4	17.8	2785	2743
	CC_{NL}	168.5	168.7	10.8	10.5	2635	2765
	CC_L	169.1	170.3	12.2	11.7	2714	2875
	CC_LNT	171.1	176.4	10.7	10.5	2557	2780
Other studies (global cropland)		115[b]; 140[c]; 157-210[d]; 164[e]		14-20[f]; 23[g]; 26[h]; 31[i]		2806[j]	

a) Summed yield of four crop types: maize, wheat, rice, and soybean. **b)** Ren et al. (2020), 0-50 cm, 1667 Mha. **c)** Zomer et al. (2017), 0-30 cm, 1631 Mha. **d)** Jobbágy & Jackson (2000), the estimate for 0-100 cm is 157 Pg C, and that for 0-200 cm is 210 Pg C, 1400 Mha. **e)** Global Soil Data Task (2014), 0-100 cm, 1518 Mha. **f)** Smil (1999). **g)** Liu et al. (2010). **h)** Lin et al. (2001). **i)** Liu et al. (2019). **j)** FAOSTAT (2023); reported total production in the year 2014 were used for comparison: 1040, 729, 731, and 306 million tonnes for maize, wheat, rice, and soybean, respectively.
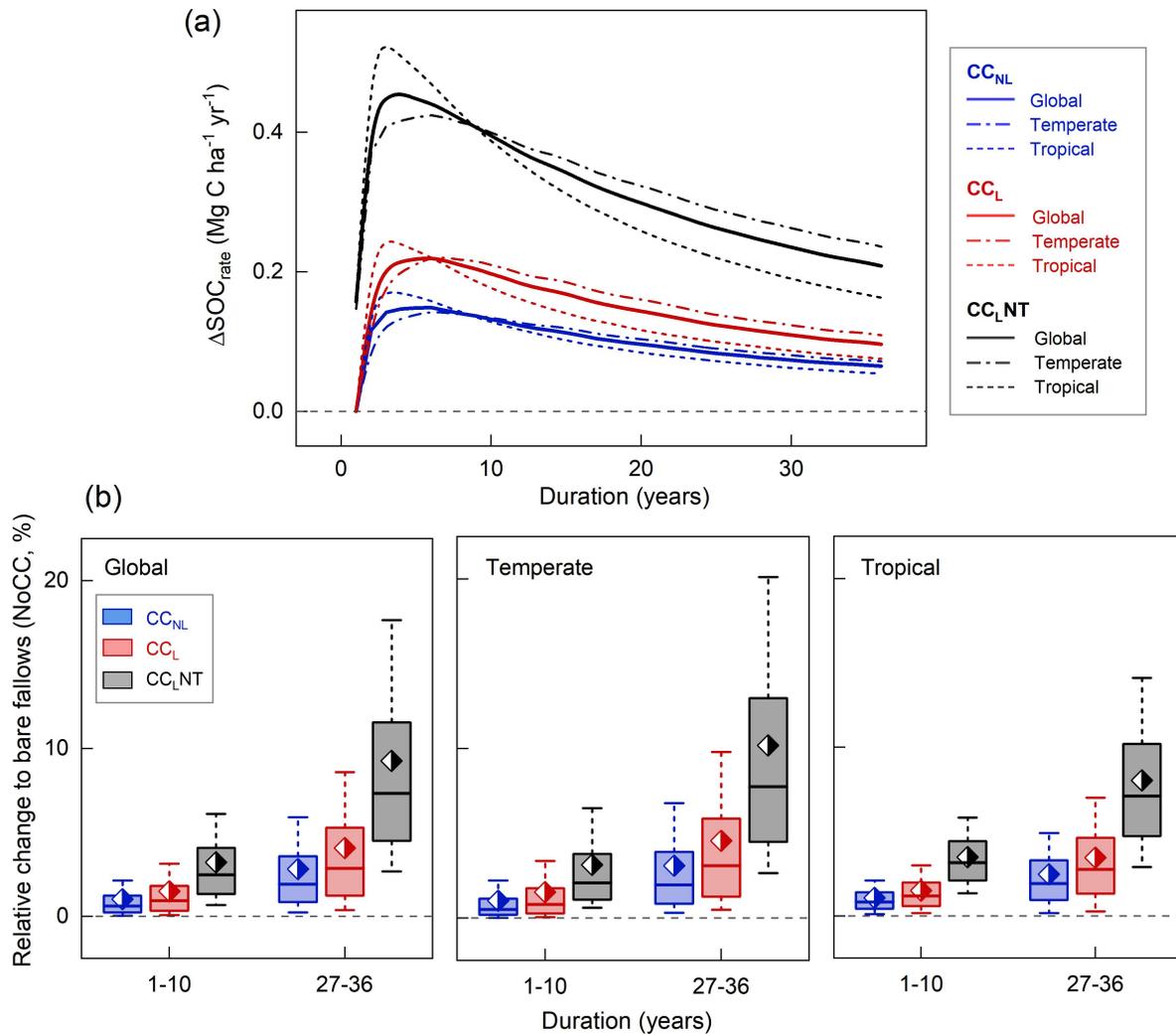
Figure 4.4 Area-weighted aggregated average annual soil C sequestration rate (Eq. 2.24; Mg C ha^{-1} yr^{-1}) across global (1597 Mha), temperate (987 Mha), and tropical (606 Mha) croplands for three cover crop practices in the CA$_{all}$ scenario (a), and relative responses (%) of SOC stocks to these cover crop strategies compared with the control treatment (bare fallows) in the first and last decades of the 36-year simulation period (b). The temperate region here is defined as the latitudes from 23.5° to 60° N/S of the equator, and latitudes between 23.5°S and 23.5°N are classified as the tropics. Box plots in (b) denote the 5th and 95th percentiles with whiskers, median and interquartile range with box lines, and mean with diamonds across all cropland grid cells (global:35039; temperate:21223; tropical:12942).

4.1.2.2 Cropland N leaching and yields

In addition to soil C benefits, cover crops resulted in a reduction in simulated N leaching in most global croplands (i.e., CA$_{all}$ scenario), with the largest decreases (~75-90%) found in Russia and large parts of Africa, regions where mineral N fertilizer application were rather low. Modelled N leaching reduction in response to CCs in China, Western Europe, and the United States—areas with

substantial fertilizer application—were still 0-45% for the 36-year simulation period (Fig. 4.5). Our simulated total nitrogen loss of 17.8-18.4 Tg yr^{-1} from fallow cropland (NoCC) was in good agreement with statistics-based estimates of 14-23 Tg N yr^{-1} (Smil, 1999; Liu et al., 2010), but lower than the findings of 26-31 Tg N yr^{-1} in Lin et al. (2001) and Liu et al. (2019) who uses a modelling approach (Table 4.2). Replacing bare fallows with cover cropping across global croplands was simulated to reduce N leaching by 7.3-7.6 and 6.1-6.2 Tg N yr^{-1} in the CC_{NL} and CC_L runs, respectively. The latter (i.e., CC_L) was ~17% lower than the decreases of 7.3-7.7 Tg N yr^{-1} from CC_LNT (Table 4.2), supporting arguments for practicing conservation tillage techniques to mitigate hydrological N losses.
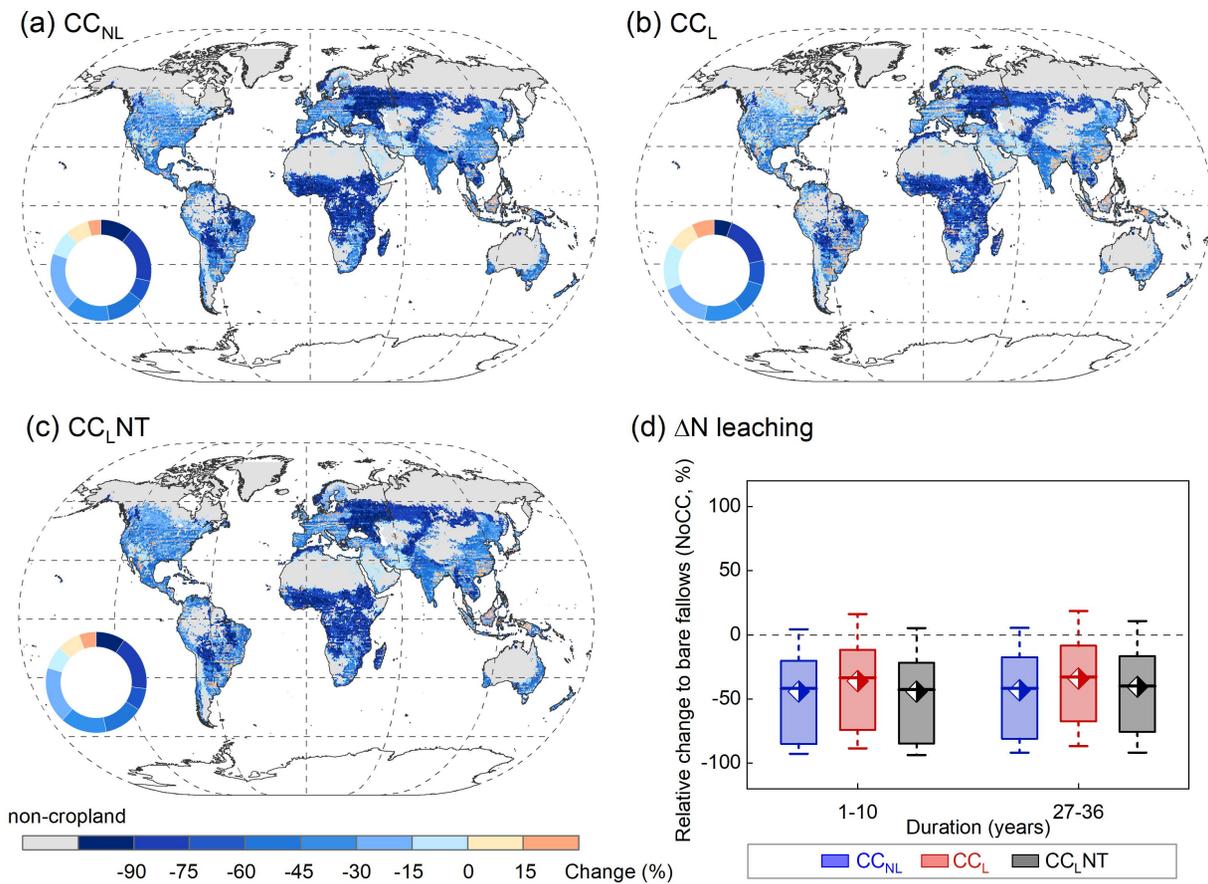
Figure 4.5 Maps of simulated responses (%; 36-year average) of cropland N leaching to CC_{NL} (a), CC_L (b), and CC_LNT (c) managements, relative to the control treatment with bare fallows (NoCC) in the CA_{all} scenario. Box plots of these responses in the first and last simulated decades are shown in (b), denoting the 5th and 95th percentiles with whiskers, median and interquartile range in box lines, and mean with diamonds across all cropland grid cells (35039). The inset donut plots represent the area proportion of each classified N leaching change from the total cropland area.
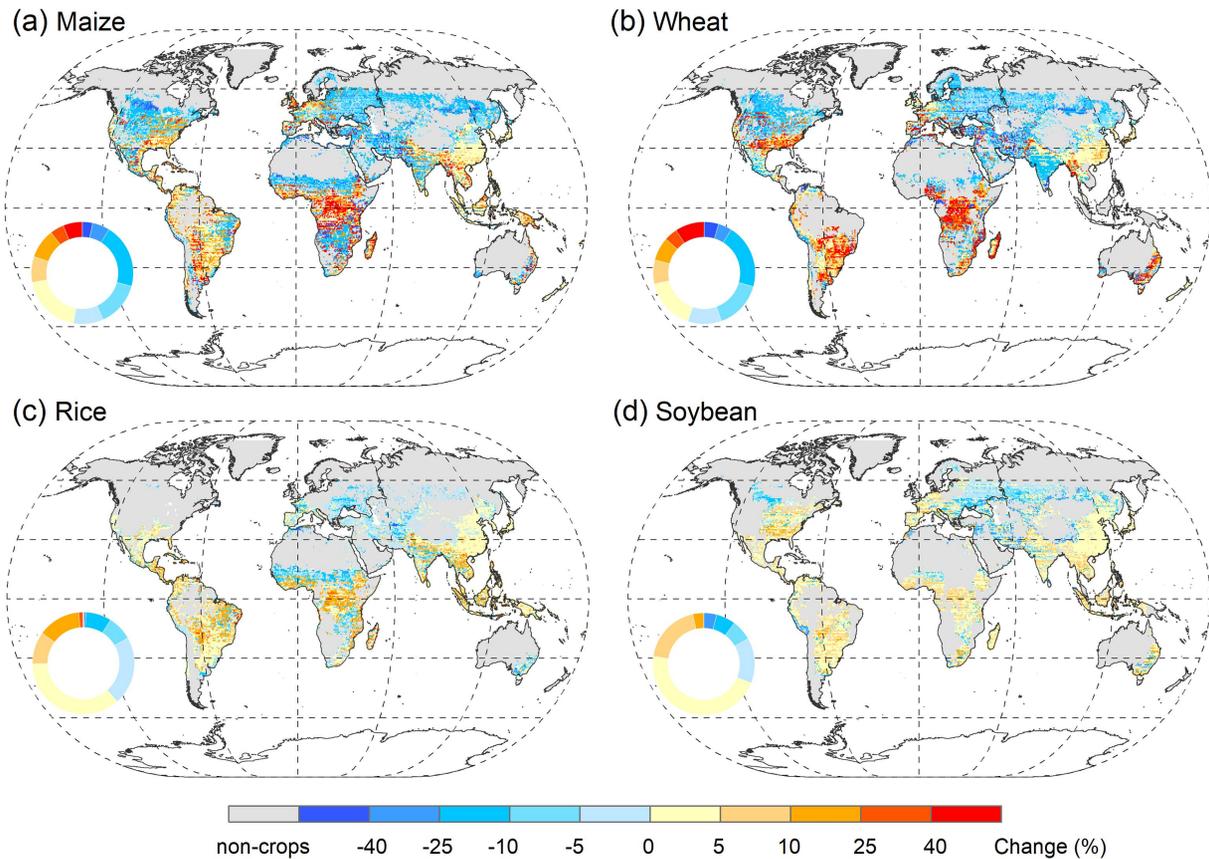
Figure 4.6 Maps of simulated response (%; 36-year average) of main-crop yield to CC_L management relative to the control treatment with bare fallows (NoCC) in the CA_{all} scenario: maize (a), wheat (b), rice (c), and soybean (d). Modelled crop yield at each grid cell is calculated as the area-weighted aggregated results in rain-fed and irrigated conditions. Global total cropping areas in 2014 modelled here are 184.5, 247.7, 151.7, and 95.9 Mha for maize, wheat, rice, and soybean, respectively, with rain-fed proportions of 84, 77, 44, and 94% for those four crop types. The inset donut plots represent the area proportion of each classified yield change from the total crop-specific area.

The modelled impacts of legume cover crops (CC_L) on yields of the main crops showed large spatial variation (Fig. 4.6). Small, and inconclusive with respect to their direction, yield changes between -5% and 5% (36-year average) were found in China across all crop types, likely as a consequence of the high N fertilizer input. A widespread yield loss in response to CCs was seen in northern cold and temperate dry climates, whereas yields in humid regions—such as the eastern USA, southern China, and most of South America and Africa—increased (Fig. 4.6), reflecting high biomass and high N fixation rates (see Sect. 4.2.3 below for details). However, these modelled impacts varied widely between different cropping systems, with the largest yield variability found

in maize and wheat, followed by rice. Productivity of soybean crops responded only little to legume CCs (Fig. 4.7).
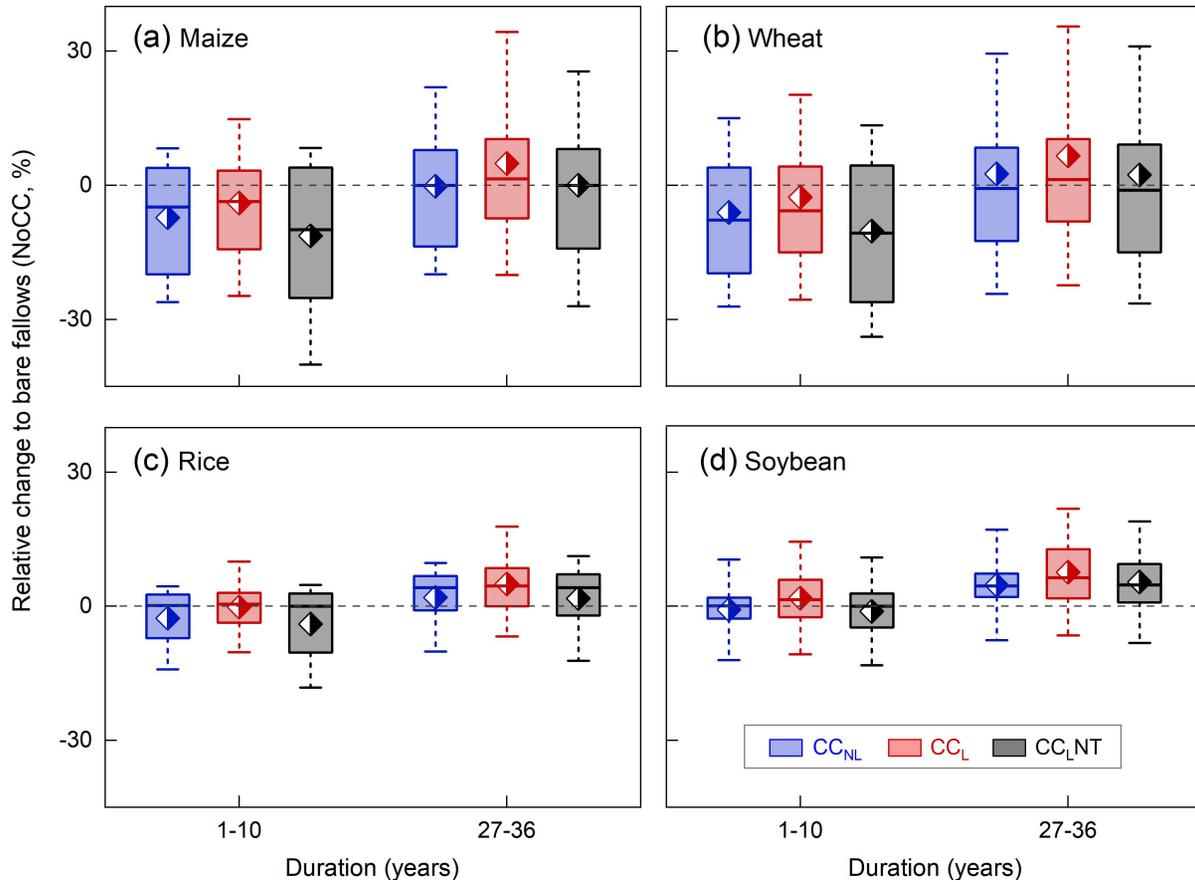
Figure 4.7 Box plots of simulated response of main-crop yield to CC_{NL}, CC_L, and CC_LNT managements, relative to the control treatment with bare fallows (NoCC) in the first and last simulated decades under the CA_{all} scenario: maize (a), wheat (b), rice (c), and soybean (d). Modelled main-crop yield at each grid cell is calculated as the area-weighted aggregated results in rain-fed and irrigated conditions. Box plots of yield relative changes (%) denote the 5th and 95th percentiles with whiskers, median and interquartile range with box lines, and mean with diamonds across all crop-specific grid cells (maize: 31635; wheat: 27126; rice: 21598; soybean: 23306).

Our model simulations under bare fallow management (NoCC) resulted in a total crop production of 2743-2785 million tonnes per year globally, consistent with FAO-reported estimate of 2806 million tonnes in the year 2014 (Table 4.2), implying the reliability of the current model version to reproduce food production at the global scale. Compared with fallow soils during off-season period, using cover crops was modelled to potentially reduce main-crop yield in the first decade for the 36-year simulation, with mean decreases of 6, 3, and 8% in CC_{NL}, CC_L, and CC_LNT, respectively. However, these negative yield effects were gradually diminished over the course of simulation, and

turned to positive impacts in the last decade, with slight production increases of 1-5% simulated for the three assessed managements in comparison with the control treatment (Table 4.2).

4.2 Discussion

4.2.1 Soil carbon stocks

LPJ-GUESS simulates cropland soil carbon stocks across all the evaluated sites well, although the measured SOC increase in response to CCs is generally underestimated (Fig. 4.1). One likely explanation for this discrepancy is the low biomass production of CCs in the model experiments (not shown), resulting in less C inputs to the soil pools compared with the field measurements. Experimental evidence from the field sites has shown that the amount of biomass C added to the soil through CCs varies widely between cover crop species (Sainju et al., 2002; Constantin et al., 2010). Using two grass functional types to represent all cover crop situations in our standardized evaluation cannot reflect this variability. Also, when comparing herbaceous CC effects on soil carbon stocks, belowground C input via roots has been proven to stably enhance SOC sequestration in the field measurements (Rasse et al., 2005; Blanco-Canqui et al., 2015). For instance, in a 2-year U.S. trial, Kuo et al. (1997) found that the root-to-shoot ratio of plant biomass C grown under natural conditions ranged from 0.5-0.8 for ryegrass (non-legume) and 0.2-0.5 for hairy vetch (legume). In comparison, higher root-to-shoot ratios in perennial grasses ranging from 1.0-3.5 were reported in another U.S. field experiment with continental climate, depending on soil sampling depth and nutrient availability (Sainju et al., 2017). In our study we implemented a prescribed root-to-shoot ratio of 2.0 to broadly represent below- and aboveground biomass productions in herbaceous plants based on literature values (see Sect. 2.1). Whether or not this set value affects the simulated root-derived carbon from CCs is difficult to assess because root biomass information was typically unavailable from the test sites.

Our modelled global-scale small SOC increase of 1.0-2.8% for non-legume cover crops (CC$_{NL}$) and 1.5-4.1% for legumes (CC$_L$) (Fig. 4.4) agreed with the meta-analysis of Poeplau & Don (2015) and Abdalla et al. (2019), in which replacing bare fallows with CCs statistically showed no significant difference between cover crop types for SOC sequestration, with a mean increase of 4.1% and 4.5% found for non-legumes and legumes, respectively. However, these reported impacts were somewhat lower than a recent synthesis conducted by Jian et al. (2020), who found that cover cropping would result in a net SOC sequestration of 0.56 Mg C ha^{-1} yr^{-1} with all cover crop types included, ~15.5% higher than the bare-fallow control treatment. In our model experiments, only the combined

agricultural practice, i.e. legume CCs and no tillage (CC_LNT), produced a mean SOC increase of 9.7% after a 36-year simulation (Fig. 4.4), which is more comparable to but still below the findings in Jian et al. (2020). The discrepancy between the global simulation and site-level field experiments likely reflects their difference in the investigated geographical scales and land-use history, as well as to the diverse managements and methodologies among field studies (such as CC species and retained residue proportion). Nevertheless, the potential of obtaining higher SOC stocks via cover crop management seems realistic, even though the exact magnitude of the effect remains unresolved.

In the global experiment, the annual SOC sequestration rate was modelled to be largest in the early years after introduction of CCs, and it then gradually declined over the course of the remaining simulation period (Fig. 4.4), similar to published findings. Sommer & Bossio (2014) reported annual SOC stock changes in response to the improved agricultural practices approaching a maximum between the third and seventh year after adopting soil-conserving techniques and a subsequent decreasing trend for 15-20 years. A meta-analysis of tropical crop ecosystems also indicated reduced SOC sequestration rates (after an initial peak) to persist for 4-25 years until a new SOC equilibrium state was reached, but the duration was highly dependent on climates and soil types (Powlson et al., 2016). In our model experiments, at the end of 36-year simulation the continued trends indicate that a new steady state in soil C and N pools had not yet been achieved, which was similar to results in Porwollik et al. (2022), who found no dynamic steady state after 50 years of simulation with the LPJmL model in response to planting herbaceous CCs in global croplands during fallow period.

In our study we attempted to quantify the contribution of CCs to enhancing soil C pools globally, which could also be interpreted as a climate change mitigation measure. After 36 years of implementation, using two herbaceous CCs was found to sequester ~0.01 Pg yr^{-1} soil carbon across the simulated 126 Mha cropland (CA_{his} scenario, ~8% of current cropland areas worldwide; Table 4.2). If all agricultural lands were to adopt cover crop practices (CA_{all} scenario), the SOC sequestration potential could be as high as 0.11, 0.15, and 0.32 Pg C yr^{-1} (i.e., 0.40, 0.55, and 1.17 Pg CO_2 yr^{-1}) for non-legumes (CC_{NL}), legumes (CC_L), and the combined agricultural practice (CC_LNT), respectively, compensating for 8-22% of annual direct greenhouse gas (GHG) emissions from crops and livestock activities (5.3 Pg CO_2eq yr^{-1}; FAO, 2020), or equivalent to 10-29% of GHG emissions from agricultural land use change (4.0 Pg CO_2eq yr^{-1}; FAO, 2020). Planting anywhere near 100% of global cropland with CCs is impractical for a number of reasons: a large

share of agricultural area used for winter crops (Poeplau & Don, 2015; Kaye & Quemada, 2017), potential water limitations or too low winter temperature during off-season periods (Dabney et al., 2001), and insufficient growing windows for CCs in multi-cropping systems in the tropics (Hu et al., 2018). Nevertheless, these estimates from our simulations do provide an upper bound for the amount of atmospheric carbon that might be sequestered through cover crop cultivation. Under the more realistic adoption scenario of CA_{pot} (590 Mha, ~37% of current cropland areas; Table 4.2), carbon sequestered by individual cover crop practice (0.15 and 0.22 Pg CO_2 yr^{-1} for CC_{NL} and CC_L, respectively) and the combined conservation management (CC_LNT; 0.46 Pg CO_2 yr^{-1}) was expected to approximately offset 3-9% of direct yearly GHG emissions from crops and livestock activities.

4.2.2 N leaching

Both model and field experiments showed that N leaching from cropland ecosystems was strongly associated with N management: applying chemical fertilizer resulted in higher hydrological N loss compared with the unfertilized treatments (Fig. 4.2), likely a consequence of the enhanced size of the nitrate pool. However, several disagreements between simulated and measured N leaching were found in some field trials despite of similar N fertilizer inputs (Fig. 4.1), indicating that other factors, such as soil texture type and drainage water, are at play as well. For example, two of the field experiments included in our analysis sites showed a decreasing trend in total N leaching (mineral plus organic) from coarse-, medium-, to fine-textured soils (Lemola & Turtola, 2000; Aronsson et al., 2011). When testing our simulation setup at these two locations, the reported soil texture effect was not captured well by the model (not shown), suggesting that the N leaching representation in LPJ-GUESS should be further improved. Moreover, compared with observations, the overall smaller reduction in N leaching in response to the simulated CCs (Fig. 4.1) might be partially attributed to the underestimated biomass of CCs, which would also underestimate plant N demand and soil N uptake. In addition, since the model cannot simulate two plants growing at the same time, the total length of the undersown-CC growing period in our simulations was approximately 1-2 months shorter than the field trials across all northern European sites (see Sect. 2.4.1), which further limited cover crop capacity for uptake of excess N remaining in the soil column in the model.

Compared with the bare-fallow setup, mean decreases of 41% and 34% in N leaching were simulated across the globe in response to the experiment with non-legume (CC_{NL}) and legume cover crops (CC_L), respectively (Table 4.2), close to the lower end of the wide reported reduction range between 30-70% in the literature (Tonitto et al., 2006; Quemada et al., 2013; Thapa et al., 2018;

Abdalla et al., 2019; Nouri et al., 2022). The reduction in N leaching due to CCs partially reflects the decreases in leachate volume and soil reactive N concentration because of enhanced water and N uptake by CCs during their growth (Thapa et al., 2018; Blanchy et al., 2023). This process may also underlie the smaller decreases in N leaching under N-fixing CCs compared with non-legumes for both field measurements (Abdalla et al., 2019; Nouri et al., 2022) and model simulations (CC_{NL} vs. CC_L). Where biological N fixation is the dominant N source for leguminous plants, it diminishes the capacity for mineral N uptake from soils (Fontaine et al., 2022). Moreover, including the no-till technique in cover cropping in our simulations had the potential to further mitigate N leaching (41% in CC_LNT vs. 34% in CC_L; Table 4.2) mainly due to the reduced net N mineralization rates (Fig. 4.8). This is in line with the findings from a meta-analysis by Thapa et al. (2018) and a recent modelling study by Porwollik et al. (2022).
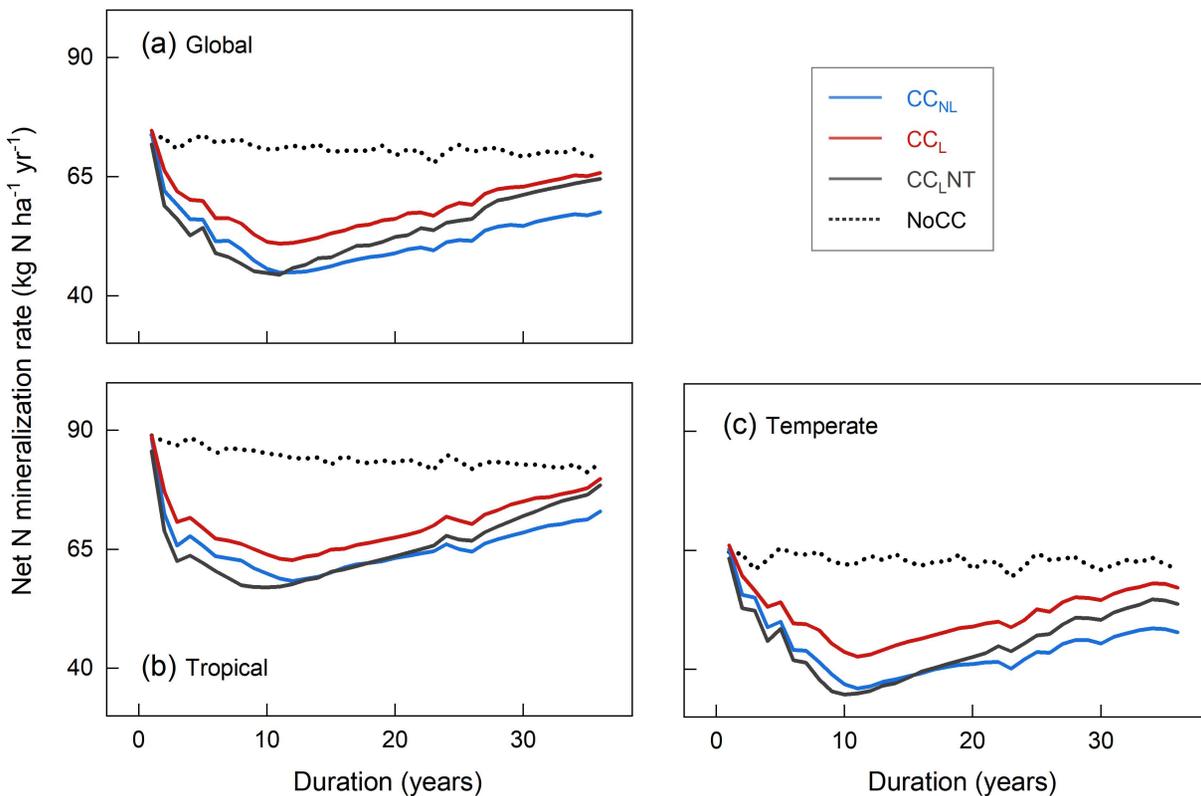
Figure 4.8 Area-weighted aggregated average annual net N mineralization rate (kg N ha^{-1} yr^{-1}) across (a) global (1597 Mha), (b) tropical (606 Mha) and (c) temperate (987 Mha) croplands with four cover crop managements (CC_{NL}: blue; CC_L: red; CC_LNT: black; NoCC: black dashed) under the CA_{all} scenario over the 36-year simulation period. Temperate region here is defined as the latitudes from 23.5° to 60° N/S of Equator, and latitudes between 23.5°S and 23.5°N are classified as the tropics.
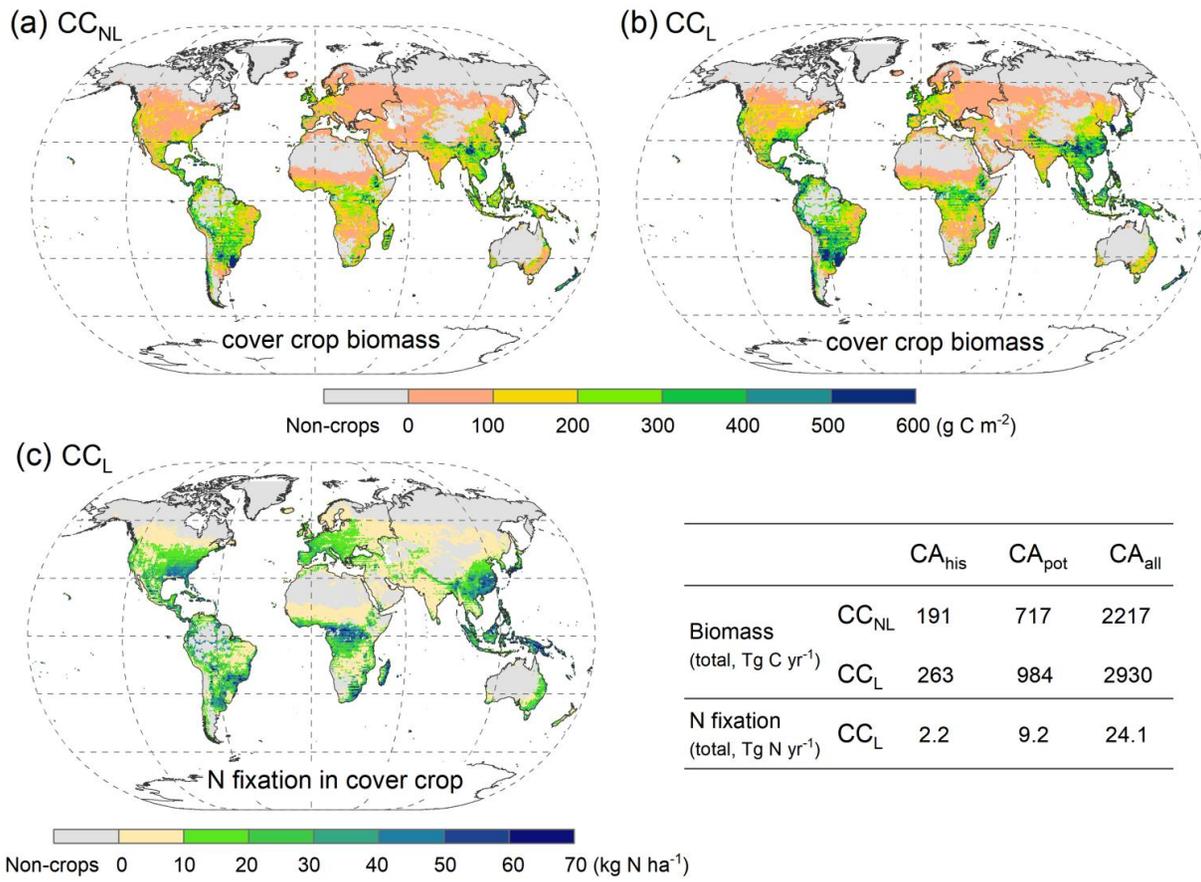
Figure 4.9 Maps of annual cover crop biomass (shoot and root) returned to the soils in the CC_{NL} (a) and CC_L (b) simulations, and biological N fixed by legume cover crop in the CC_L run (c). Biomass and N fixation rates in maps are given as the mean values for the 36-year simulation period. See Sect. 2.4.2 for details on three conservation agriculture (CA) area scenarios.

Globally, the largest percent decreases in N leaching due to CCs were modelled in regions with relatively little N fertilizer use (such as Russia and large parts of Africa), where soil reactive N pools were small. Results from a six-year field experiment implemented by Wittwer et al. (2017) also showed that the effectiveness of CCs in reducing N leaching decreased with management intensity (e.g., tillage regimes and fertilizer application rates). This effect underlies discrepancies at some national borders, such as Indonesia and Papua New Guinea (Fig. 4.5), countries with similar climates but with contrasting fertilizer applications. Likewise, in some arid and semi-arid regions, as well as temperature-limited areas in the high latitudes (e.g., Canada) a slight decrease of N leaching in response to cover cropping systems was found, as poor growth conditions constrained the CC capacity for soil N uptake. In addition, the rapid turnover rate of SOM pools driven by warm and moist climate (Olin et al., 2015a), together with abundant precipitation may increase N leaching

with cover crop practices in the humid tropics as a result of high biomass of N returned to soils (Fig. 4.9) and enhanced water drainage (Porwollik et al., 2022).

4.2.3 Crop yields

Accounting for the impacts of management practices, particularly regarding water and N limitations to crop growth in LPJ-GUESS, resulted in a good agreement between simulated and observed crop yields across different field trials. For both modelling and field-based experiments, yields in the main crops following non-legume CCs declined, although overall difference to fallow controls (NoCC) was small (Table 4.1). The difference between periods of soil N mineralization and high N demand of main crops (Marcillo & Miguez, 2017), and enhanced soil N immobilization shortly after the planting of non-legume CCs (Erenstein, 2003; Abdalla et al., 2019) may contribute to the declines in yields of the main crops in the field experiments. In comparison, N-fixing CCs with relatively low C:N ratios are expected to stimulate soil N release during their decomposition, enhancing plant available N in soils (Quemada et al., 2013; Thapa et al., 2018). This was in line with our model findings, wherein legume CCs generally resulted in higher net N mineralization rates than non-legumes (Fig. 4.8), and thus increased the productivity of the main crops in some cases (Fig. 4.1). However, it should be noted that these CC effects were highly dependent on cropping systems, with little impacts found on productivity of soybeans (Table 4.1). This is likely due to their N fixation capacity, which diminished the N competition between CCs and soybeans in both field trials and model simulations.

Our modelled global mean main-crop yield losses due to CCs in the first decade of the simulations (-3% for CC_L and -6% for CC_{NL}; CA_{all} scenario in Table 4.2) compared well with a recent meta-analysis by Garba et al. (2022), who reported a mean crop production change of -4.9% and -10.1% for legume and non-legume CCs, respectively, after 2-17 years of management. Main-crop yield reduction under cover cropping systems likely reflected (a) the indirect competition for water and nutrients between CCs and subsequent main crops (Valkama et al., 2015), and (b) the time that soil SOM pools need to adjust to management shifts (Figs. 4.4 and 4.8). Garba et al. (2022) also pointed out that cover cropping systems under the no-till practice resulted in lower main-crop yields compared with conventional tillage, in line with our model findings in terms of total crop productions worldwide (CC_LNT vs. CC_L; Table 4.2). However, at least in our simulations, these negative yield effects induced by conservation tillage may be mitigated over the course of the simulation (Table 4.2) because of the gradual stabilization of soil C and N pools over time (Figs. 4.4

and 4.8). A similar finding from a meta-analysis by Pittelkow et al. (2015) indicated that yield benefits, globally, in cereal- and legume-based cropping systems may be attained after 10+ years of conversion from conventional tillage to no-till management.

N fertilizer application was found to be another factor that influenced the effectiveness of CCs on subsequent crop yields for both site-level (Fig. 4.2) and large-scale simulations (Fig. 4.6). The smallest impacts on main-crop production were found for well-fertilized cover cropping systems, consistent with previous field-based reviews (Tonitto et al., 2006; Quemada et al., 2013; Marcillo & Miguez, 2017; Daryanto et al., 2018; Zhao et al., 2022), since enhanced soil mineral N pools driven by fertilization reduce the N competition between CCs and main crops. This can explain the small yield penalty (or benefit) from cover cropping in soybean (Figs. 4.6-4.7), which is a nitrogen fixer and experiences less N stress during the growing season compared with cereal crops. Likewise, the spatial variability regarding CC impacts on rice production was also much smaller than simulated maize and wheat CFTs (Figs. 4.6-4.7), primarily because rice in our simulations was mostly irrigated (Fig. 4.6), which reduced water limitation on crop growth caused by CCs in rice-producing areas. Furthermore, the broadly negative impacts of CCs on simulated yields in northern temperate climatic regions (Fig. 4.6) can be attributed to the slow decomposition of SOM in response to low temperature, where the N retained in the SOM is released evenly throughout the year and not easily available for main crop uptake after CC growth (Olin et al., 2015a). In contrast and as discussed above, plant materials from CCs in the humid tropics are expected to rapidly decompose due to the fast turnover rate, continuously releasing reactive N for plant uptake in the next cropping season and therefore enhancing main-crop productions. This contrasting spatial difference in yield changes between temperate and tropical climates supports a meta-analysis finding that cultivating CCs during bare-fallow period, on average, has a risk to reduce main-crop productivity by ~12% in temperate agricultural soils while gaining ~15% of yield benefits in the tropics (Garba et al., 2022).

4.3 Conclusions

In this chapter we investigated the influence of cover cropping management on global crop ecosystems using the updated LPJ-GUESS model version from Chapter 3. The simulated C-N variables and main-crop productions in response to two herbaceous cover crop types (i.e., non-legumes and legumes) were widely evaluated against measured data from site level to global. Our model estimates revealed that crop ecosystems implemented in LPJ-GUESS realistically responded

to non-legume and legume cover cropping under a range of water and N managements, and resulted in comparable C-N variables with observations, particularly for cropland SOC stocks.

Our simulations demonstrated that the impacts of CCs on global agricultural lands can be beneficial for long-term environmental sustainability without compromising crop productions, particularly for the integrated management practice with N-fixing CCs and no-till technique included. This combined strategy was modelled to potentially increase soil carbon levels by 7% (+0.32 Pg C yr^{-1}) and reduce cropland N leaching by 41% globally (-7.3 Tg N yr^{-1}; 36-year average), but with a 8% of loss in total crop production during the first simulated decade. However, these negative effects on crop production were diminished over the course of managements and turned to a slight increase (~ 2%; +37 million tonnes per year including maize, wheat, rice, and soybean) after 36 years of implementation. In comparison with non-legume CCs, the adoption of N-fixing cover cropping, at least in our model experiments, contributed more to gaining yield benefits in humid tropics while it is mitigating production losses under northern temperate climates. This spatial variation due to CCs was also found to be associated with main-crop types and N fertilizer inputs, with little yield changes simulated in soybean systems and highly fertilized agricultural soils.

The dynamic process of N fixation for grass CCs in LPJ-GUESS provides an opportunity to overall assess atmospheric carbon and nitrogen flows to agricultural lands during fallow periods, and thus is relevant for the estimates of global terrestrial C-N fluxes and pools under present-day and future climate. It can also help to examine the possibility of conservation agriculture for achieving sustainable food production, through comparing with the impacts from other management strategies.

5 Impacts of agricultural management practices on soil carbon stocks, nitrogen loss, and crop production in eastern Africa

This chapter further investigates the potential responses of crop ecosystems to improved agricultural practices, in addition to cover cropping management shown in Chapter 4. Here, we choose two eastern Africa countries—Kenya and Ethiopia—as a regional case study, primarily due to their severe land degradation on agricultural soils over the past decades. The objective of this section is to examine whether these alternative management practices favor soil degradation restoration and crop production while minimizing N loss (including gaseous emissions and N leaching) from the cropland situated in the tropical climates. Model results are first assessed on experimental data from two long-term (>10 years) field sites in western Kenya and compared against country-level yield statistics. The management effects on soil C pools, crop yields, and N losses in Kenya and Ethiopia are subsequently investigated under present and future climate scenarios.

5.1 Results

5.1.1 Model performance at site scale

In most cases, the simulated maize yields in the long rainy season (from March to August) were higher than the measurements, with the mean overestimation ranging from 18% to 21% at the two experimental sites (Fig. 5.1). The averaged yields over the entire experimental period (2004-2015) between simulations and observations compared well across all the evaluated treatments, with the simulated values falling within the range of measured standard deviation (Fig. 5.2). However, LPJ-GUESS did not well predict the interannual variations of the yields, producing a low Pearson correlation coefficient (r) and high mean absolute error (MAE) in all the INM3 and CT1 experiments. As expected, the measured and simulated estimates from the combined conservation managements (e.g., manure with residue retention at INM3) were higher than the individual ones in the little-fertilized treatments, but yield discrepancies between managements became small and insignificant when maize received a high N application rate of 90 kg N ha^{-1} (Fig. 5.2).

The simulated SOC at both sites showed a declining trend from 2004-2015 under all the assessed treatments, agreeing well with the observation of soil carbon loss over the same period; however, the model generally underestimated SOC at the beginning of experiment while overestimating soil carbon stocks in the last two sampling years (Fig. 5.2). A linear correlation (p<0.01) between the simulated and measured SOC stocks was found when all the managements were included, with the

model explaining 82% and 64 % of the variation in observed SOC at INM3 and CT1, respectively (Fig. 5.1). Low mean absolute error of 4.2% and RMSE value of 4.1 Mg C ha^{-1} were found for the INM3 treatments, and 3.5% and 3.9 Mg C ha^{-1} for the CT1 treatments (Fig. 5.1). The field-measurements showed that SOC stocks from the combined conservation managements were significantly higher than the conventional ones (i.e., Nx_NoMan_NoRR at INM3 and Nx_Till_NoRR at CT1). The model can broadly capture this response well, but it had difficulty in reproducing SOC difference between the individual managements (Figs. 5.1-5.2).
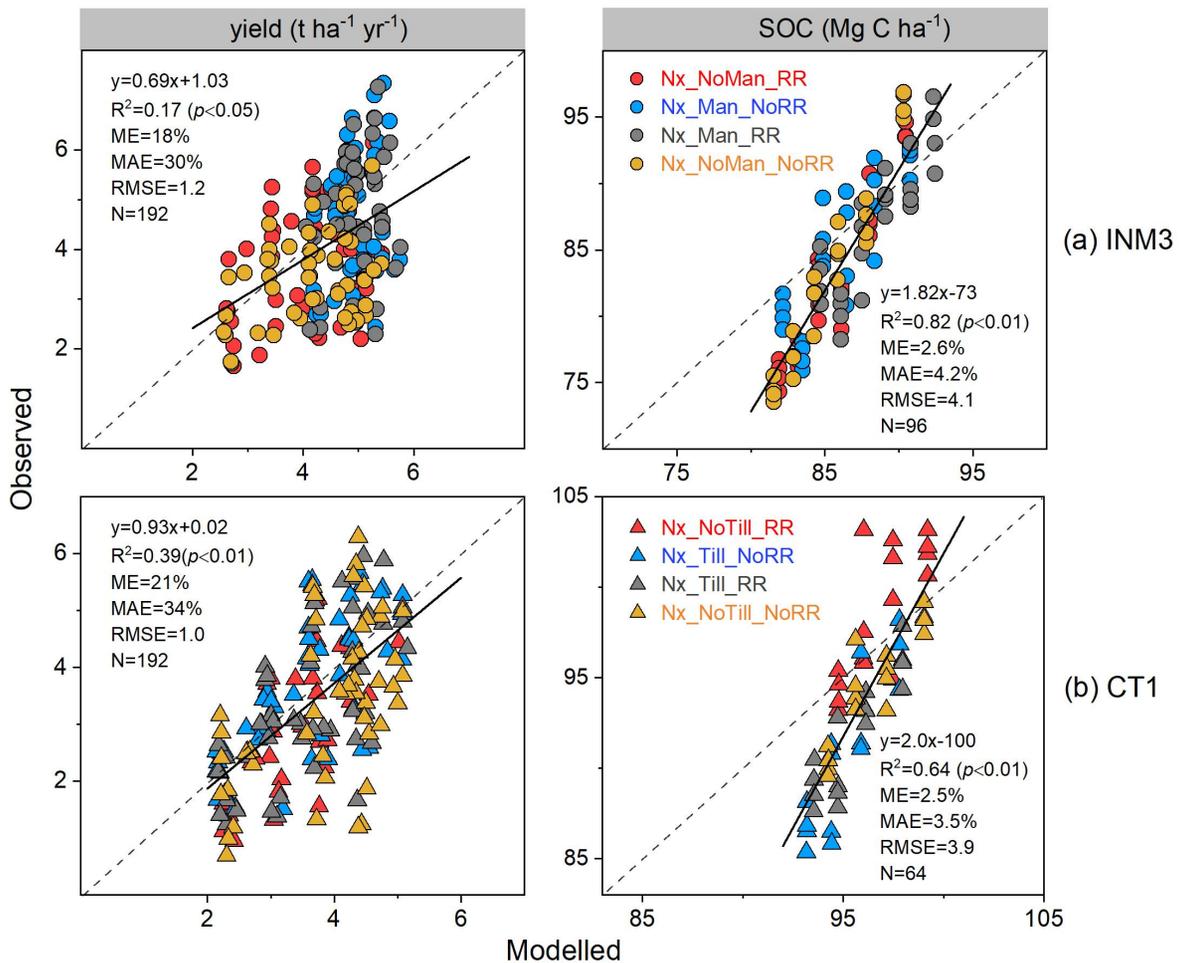
Figure 5.1 Comparison of modelled and observed maize yields (long rainy season, i.e. the main growing period) and SOC stocks (0-150 cm) at INM3 (a) and CT1 (b) sites across all evaluated treatments. The dashed line is 1:1 line and black bold line is fitted linear regression; ME and MAE are mean error and mean absolute error, respectively, representing in percent (%); RMSE is root mean square error, with the unit of t ha^{-1} yr^{-1} for yield and t C ha^{-1} for SOC. See Table 2.6 for the treatment abbreviations and their explanations.
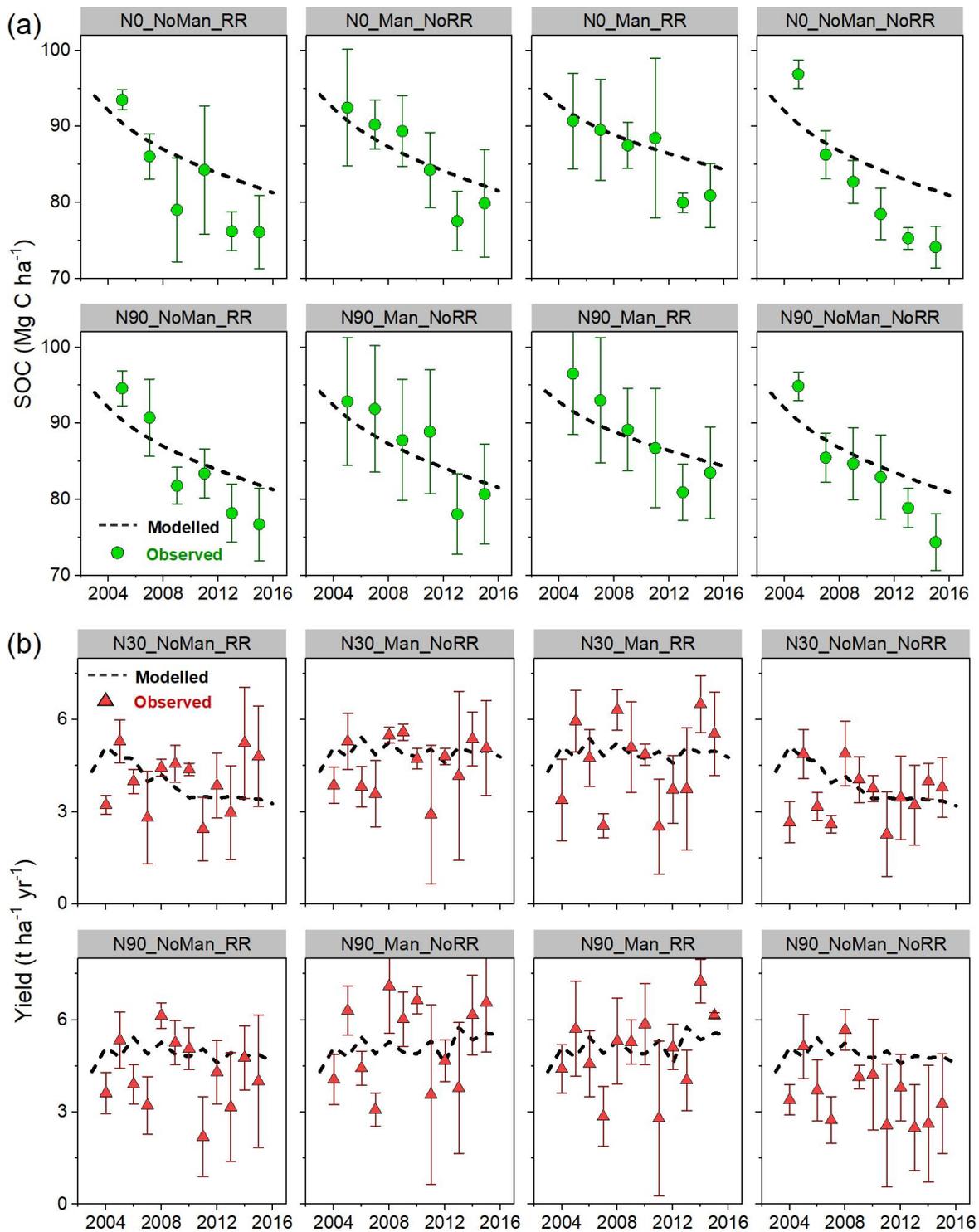
Figure 5.2 The modelled and observed SOC stocks (a) and maize yields (b) for the evaluated treatments at the INM3 site, with two levels of N fertilizer shown in plots (N0 and N90 for SOC; N30 and N90 for yields). The dashed lines are simulations, and closed circles and triangles represent the observations averaged over the four replicates in the trials, with standard deviation given in the vertical bar. See Table 2.6 for the treatment abbreviations.

Compared with observations, LPJ-GUESS in general underestimated absolute SOC loss—i.e., SOC stocks in 2015 minus the values in 2005—across all the INM3 experiments (Fig. 5.2). Due to the additional C input to soils from manure and residue retention, the simulated combination of these two managements produced the lowest loss of 6.7 Mg C ha^{-1} at INM3 site (N0_Man_RR), with this loss mitigated by N fertilizer use (6.3 Mg C ha^{-1}, N90_Man_RR). Conversely, simulation from LPJ-GUESS yielded the highest C reduction of 8.9 Mg C ha^{-1} in the unfertilized maize cropping system with no manure and no residue addition (N0_NoMan_NoRR). This estimate was somewhat close to the modelled SOC reduction of 8.6 t C ha^{-1} in maize residue only (N0_NoMan_RR) and manure application only (N0_Man_NoRR) managements. Taken together, practicing manure and residue retention strategies, combined with 90 kg N ha^{-1} of fertilizer use was simulate to mitigate SOC loss by 29% in comparison with the control treatment, lower than the observed reduction of 43% (Fig. 5.2).

5.1.2 Regional yields comparison

We simulated six crop types under the conventional management (STD; see Table 2.7 for management details) to realistically represent agricultural production in eastern Africa from 1901-2014 (referred to as 5-B1 run in Table 2.5). The simulated results from 1961-2014 and 1993-2014 were chosen to compare with annual FAO-based yield data in Kenya and Ethiopia, respectively, due to their different time frames reported in statistics.

Modelled maize yields in two countries showed a good agreement with observations, with a mean error (ME) of -6 % and RMSE value of 0.22 t ha^{-1} yr^{-1} in Kenya, and -21% and 0.54 t ha^{-1} yr^{-1} in Ethiopia (Fig. 5.3). LPJ-GUESS tended to broadly overestimate the reported yields in pulses and sorghum, with the country-level overestimation spanning from 48-257% and 72-203%, respectively. Apart from sorghum production in Kenya, the simulated and reported yields in most crop types showed a strong correlation, with a high range of Pearson correlation coefficient (r) from 0.55-0.90 (p<0.001; Fig. 5.3), reflecting that the model was able to capture the interannual variability in yields despite some disagreement between simulations and observations for individual years. Additionally, the simulated total maize production in eastern Africa increased from 5.4 million tonnes in 1993 to 9.8 million tonnes in 2014, close to the reported range of 3.6-11.2 million tonnes per year over the same period (Fig. 5.3). With all six agricultural crops included, LPJ-GUESS generated a total production of 19.7 million tonnes per year averaged over 1993-2014, ~45% above the reported value of 13.6 million tonnes per year from FAO. This overestimation in total agricultural production

was most likely due to the largely overestimated crop yields in pulses systems, where modelled yields were 4.4 million tonnes per year greater than FAO records (6.6 and 2.2 million tonnes per year for simulation and FAO statistics, respectively).
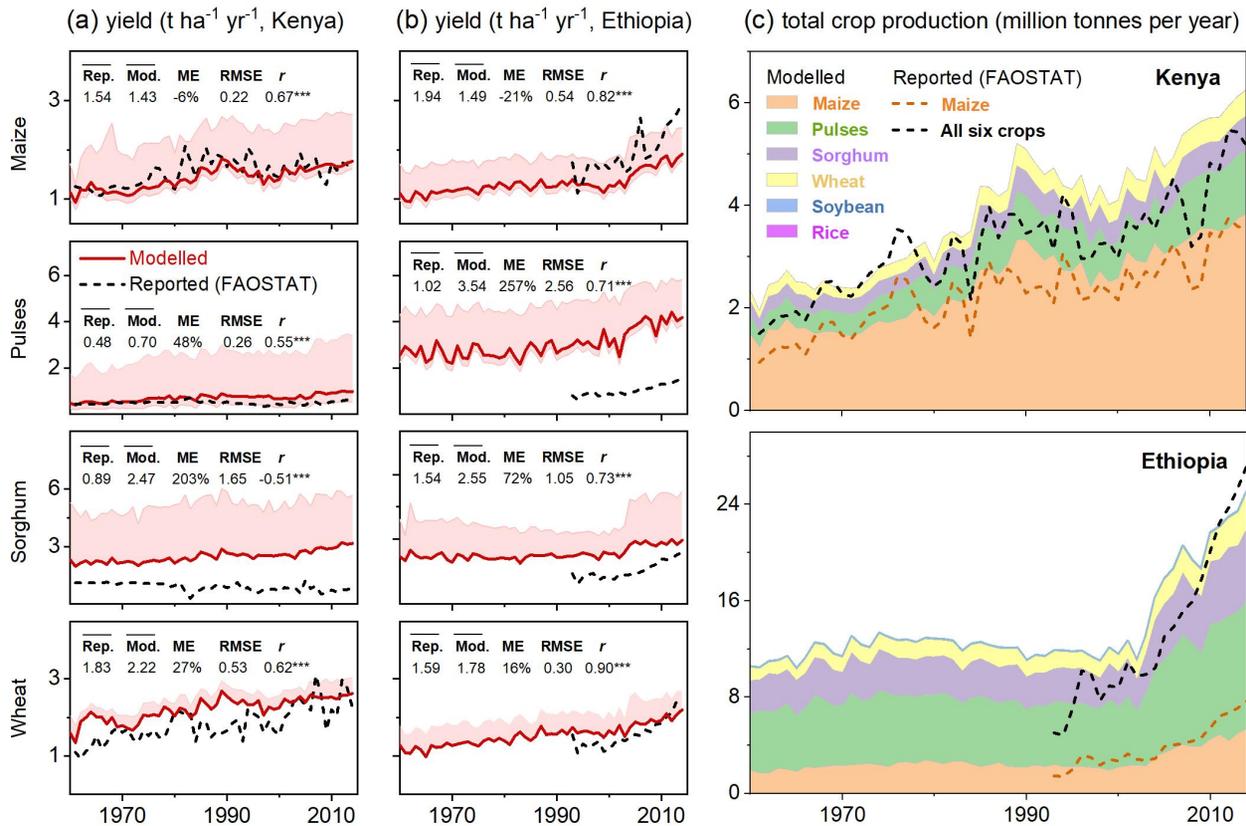
Figure 5.3 Comparison of modelled and FAO-reported crop annual yields on country level from 1961-2014 in Kenya (a), Ethiopia (b), and total crop production (c). The upper and lower bounds of shade areas in (a) and (b) represent the simulated yields under irrigated and rain-fed conditions, respectively, with their area-weighted aggregated results as given in red solid lines. $\overline{\text{Rep.}}$ and $\overline{\text{Mod.}}$ indicate the reported and simulated yields averaged over FAO-based periods (1961-2014 for Kenya and 1993-2014 for Ethiopia), respectively. ME is mean error, represented in percent (%). RMSE is root mean square error, with the unit of t ha^{-1} yr^{-1} for yields. r is Pearson correlation coefficient, where *** denote the correlation to be statistically significant at $p=0.001$ level.

5.1.3 Ecosystem responses to management practices

5.1.3.1 Historical period

Including six crop types, all the assessed management practices that address aspects of sustainable land management led to a net increase in simulated cropland soil carbon in Kenya and Ethiopia compared with the conventional management prevalent in eastern Africa (Fig. 5.4). Our simulation of the integrated conservation agriculture practice, as expected, produced the largest increase in soil

carbon sequestration of ~11%, followed by cover cropping implementation (legumes and non-legumes), residue management, and manure application, with the lowest increase of ~2% found in no-tillage management option. Most of these explored practices also gained the extra benefit of increased yields—despite being in model simulations accompanied by larger N losses (gaseous emissions and N leaching)—with the exception of cover crops in some regions. Compared with the implementation of non-legume cover crops, practicing legume cover cropping technique was simulated to generally result in larger N loss over Kenya and Ethiopia. However, this strategy was accompanied by an enhancement in modelled crop production of ~18%, most likely due to extra N addition into soils through symbiotic N fixation in herbaceous legumes, which facilitates a N-rich soil environment to subsequent crops for better growth and productivity. Returning 100% of crop residues to the field after harvest and using manure as fertilizer were the two simulations that increased crop production for most parts of eastern Africa but with the large environmental "cost" of an increase in N loss. The enhancement in both yield and N loss from residue retention might reflect that N becomes available for plant uptake over a longer period, and nothing grows during crop fallow period which can increase the N leaching from soils. In addition, no-tillage, as an important component in conservation agriculture in the tropics of Africa, had a potential (at least in our simulations) to reduce the N loss from cropland with slight yield benefits depending on region (Fig. 5.4).

The influence of individual (and integrated) management options varied widely between different parts of Kenya and Ethiopia, depending on soil properties and climate at a location, as well as on specific cropping systems implemented in agricultural soils. With all agricultural crops included, legume cover cropping practice was in general identified as a promising option for potentially sequestering soil carbon, with 43% of cropland grid cells having this technique as the optimal C management (OPT), followed by manure application (MAN), residue retention (RR), and the conventional management practice (STD, Fig. 5.4). However, this spatial pattern showed distinct difference among specific cropping systems. For instance, leaving all crop residue in the field was simulated to dominate SOC enhancement in maize and sorghum systems in eastern Africa, but it only slightly contributed to sequestering soil carbon in wheat and pulses cropping systems (Fig. 5.5), likely reflecting the differences in biomass production, phenological responses to climate change, manure application, and N fertilizer investment among these crop types.
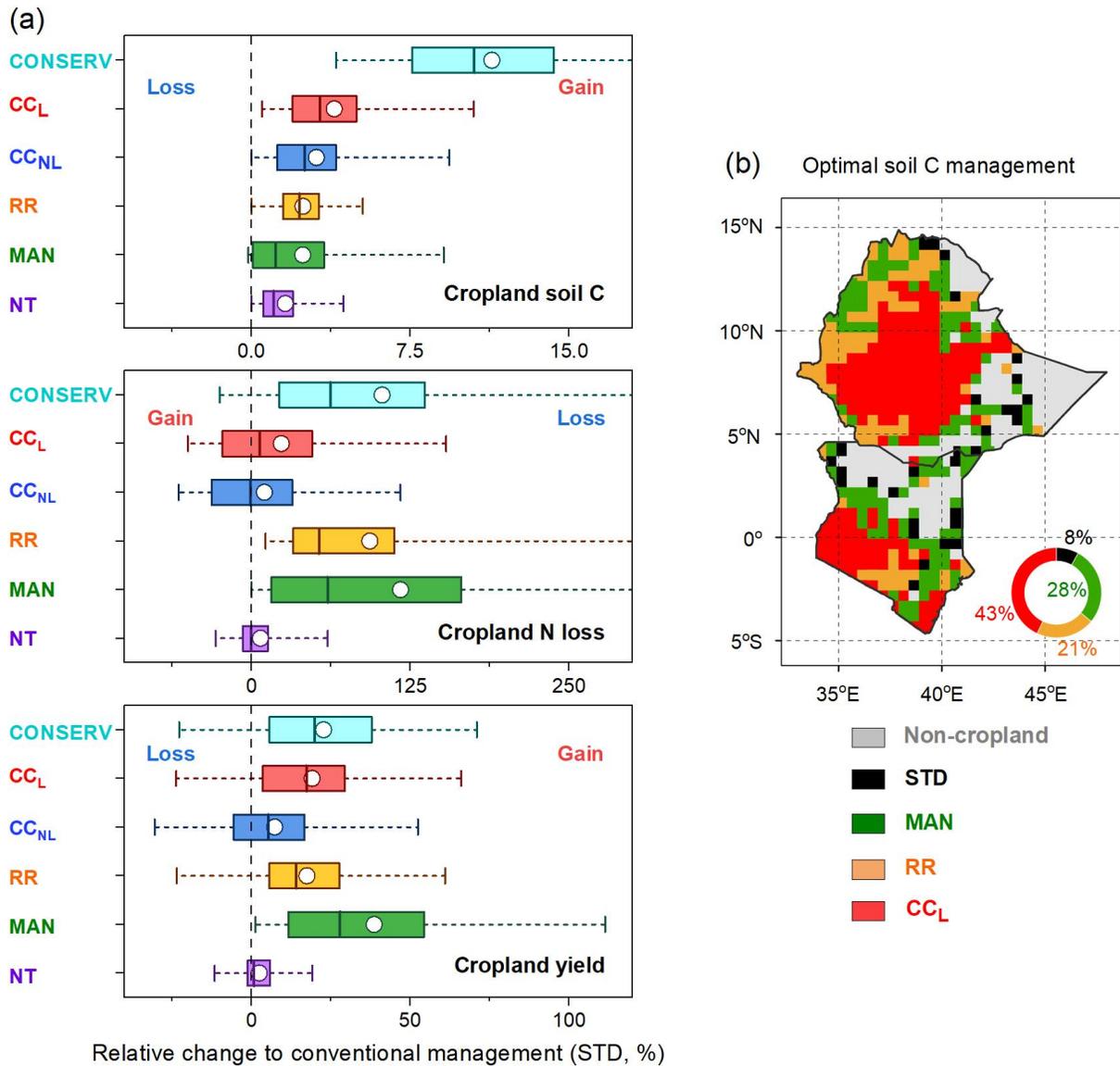
Figure 5.4 The modelled relative response (%) of cropland SOC, N loss, and yield to alternative management practices compared with the conventional management prevalent in eastern Africa (a), and the optimal SOC sequestration strategy (OPT; Eq. 2.25) simulated by LPJ-GUESS in agricultural soils over Kenya and Ethiopia (b). Box plots in (b) denote the 5th and 95th percentiles by the whiskers, median and interquartile range are the box lines, and means are symbolized as white circles across all cropland grid cells (428). The inset donut plot represents the area proportion of each optimal management from the total grid area. The conventional management (STD, black in b) was chosen when none of other alternative practices produced a net increase in SOC. Abbreviations: CONSERV – conservation agriculture; CC_L – legume cover crop; CC_{NL} – non-legume cover crop; RR – residue retention; MAN – manure application; NT – no-tillage; STD – conventional management.
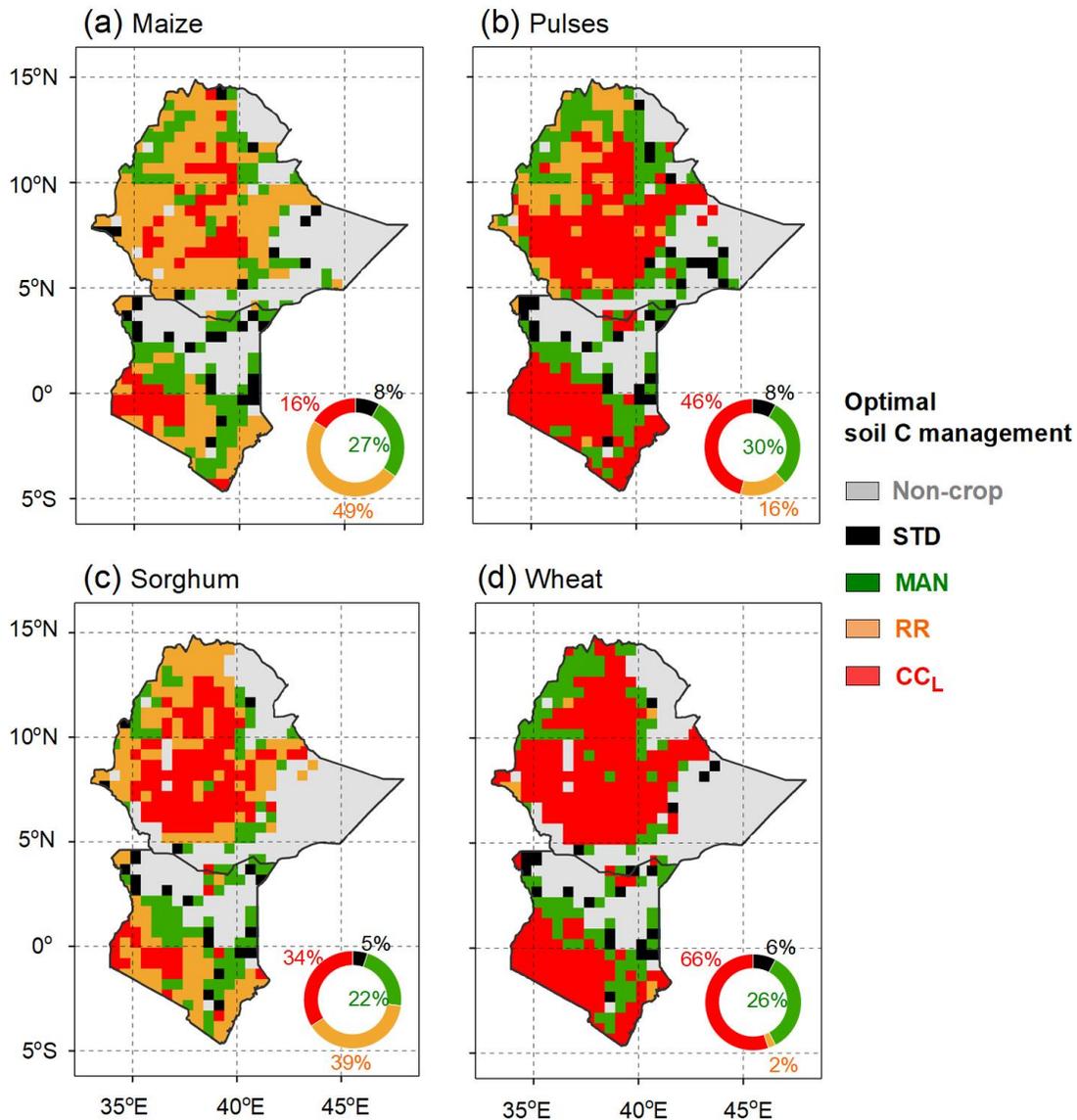
Figure 5.5 Maps of the optimal SOC sequestration strategy (OPT; Eq. 2.25) simulated by LPJ-GUESS in maize (a), pulses (b), sorghum (c), and wheat (d) cropping systems over Kenya and Ethiopia. The inset donut plots represent the area proportion of each optimal management from the total grid area. The conventional management (STD, black in plots) was chosen when none of other alternative practices produced a net increase in SOC. See Fig. 5.4 for management abbreviations.

The modelled total cropland SOC storage (0-150 cm) from different managements ranged from 932-1038 Tg C in Kenya and 2569-2895 Tg C in Ethiopia, which, as expected, was higher than the published estimates for the depth layer 0-30 cm (Zomer et al., 2017). However, these simulated soil C stocks were close to the scaled-up published values using the depth distribution functions (Eqs. 2.22-2.23), with 727-2227 Tg carbon estimated for the depth of 0-150 cm in those two counties (Table 5.1). In Kenya the simulated N loss of 45-134 Gg N yr^{-1} (1 Gg= 10^9 grams) from various

management options was comparable with the statistic-based estimates of 111 Gg N yr^{-1} in Zhang et al. (2021). Furthermore, the simulated total maize production of 8.7-14.3 million tonnes per year in eastern Africa were consistent with the FAO-reported yield of 11.2 million tonnes per year (Table 5.1). With six crop types included, an overall overestimation of 7-47% was found, primarily due to the overestimated yields in pulses and sorghum shown in Fig. 5.3.

Table 5.1 Modelled total cropland soil C stocks (0-150 cm), N loss (gaseous emissions and N leaching), and total crop production with alternative management practices in Kenya and Ethiopia, compared with literature-reported estimates. See Fig. 5.4 for management abbreviations.

Management	Soil C stock, total (Tg C)		N loss, total (Gg N yr^{-1})		Crop production (million tonnes per year)			
	Kenya	Ethiopia	Kenya	Ethiopia	Kenya		Ethiopia	
					Maize	All crops[a]	Maize	All crops[a]
STD	939	2592	61	157	3.9	7.3	7.3	21.7
NT	948	2623	64	164	3.8	7.6	7.1	23.4
MAN	932	2569	45	79	3.2	6.7	5.5	19.9
RR	957	2653	134	359	4.2	8.4	8.2	26.7
CC$_{NL}$	969	2696	64	190	3.9	7.7	7.7	23.1
CC$_L$	979	2710	75	204	4.9	8.9	8.5	24.7
OPT	993	2786	81	229	4.6	9.0	8.3	25.8
CONSERV	1038	2895	127	375	5.1	9.4	9.2	27.0
Other studies	414[b] 727[c]	1268[b] 2227[c]	111 (76-297)[d]	—	3.5[e]	5.2[e]	7.7[e]	19.6[e]

a) Summed yield of six crop types: maize, pulses, sorghum, wheat, rice, and soybean; **b)** Zomer et al. (2017); **c)** Zomer et al. (2017), soil C stocks were scaled up to 0-150 cm from the original depth of 0-30 cm using the depth distribution functions (see Eqs. 2.22-2.23); **d)** Zhang et al. (2021), the mean estimate over 2006-2015 was chosen, with a range given in bracket; **e)** FAOSTAT (2023), the reported total production in the year 2014 were used for comparison, since the simulated cropland area was fixed from 2014 onwards (~6.2 and 17.4 Mha for Kenya and Ethiopia, respectively), see 5-B2 runs in Table 2.5.

5.1.3.2 Future projection

Comparing with the standard model simulation (STD), all management practices were simulated to increase cropland SOC storage in the last decade of this century (i.e., 2091-2100) but with insignificant discrepancies among three future climate and CO$_2$ scenarios (Fig. 5.6). Although no-

tillage showed slight impacts on crop production, it was accompanied by a reduction of N loss (Fig. 5.6). A clear yield difference among three SSP scenarios was consistently seen in practices of legume cover crop and the integrated conservation agriculture, with yield increases being higher for SSP5-85 than for SSP1-26 climate pathway (Fig. 5.6). This likely suggests the stronger CO_2 fertilization impact on the growth of herbaceous legumes under SSP5-85. Overall, the future projection revealed that legume cover cropping represented a near win-win situation in terms of soil carbon sequestration and yield benefits in eastern Africa, also with lower N loss compared to manure and residue managements.

Table 5.2 The relative (%) number of cropland grid cell and areas regarding the potential transition of optimal SOC sequestration practice (OPT; Eq. 2.25), comparing the historical period (GCM-based climate; 5-C3 runs in Table 2.5) with three future SSP scenarios from 2091-2100 (5-C2 runs in Table 2.5). See Fig. 5.4 for management abbreviations.

From (GCMs historical)	To (future)	Amount of cropland grid cells (%)			Amount of cropland areas (%)		
		SSP1-26	SSP3-70	SSP5-85	SSP1-26	SSP3-70	SSP5-85
STD	RR	0.70	1.17	1.17	0.00	0.01	0.01
	MAN	2.80	3.04	2.57	0.03	0.03	0.02
	CC_L	1.87	1.17	1.17	0.03	0.03	0.03
RR	STD	0.23	0.47	0.23	0.00	0.49	0.00
	MAN	0.47	0.00	0.23	0.04	0.00	0.01
	CC_L	12.85	14.02	16.12	17.74	17.73	19.40
MAN	STD	1.64	1.87	2.34	0.01	0.01	0.02
	RR	3.27	4.21	4.21	1.12	0.38	0.50
	CC_L	14.02	15.65	15.89	9.88	11.10	11.12
CC_L	STD	0.00	0.00	0.00	0.00	0.00	0.00
	MAN	0.47	0.00	0.00	0.01	0.00	0.00
	RR	1.17	2.10	1.87	0.84	1.52	1.46
Total change		39.5	43.7	45.8	29.7	31.3	32.6
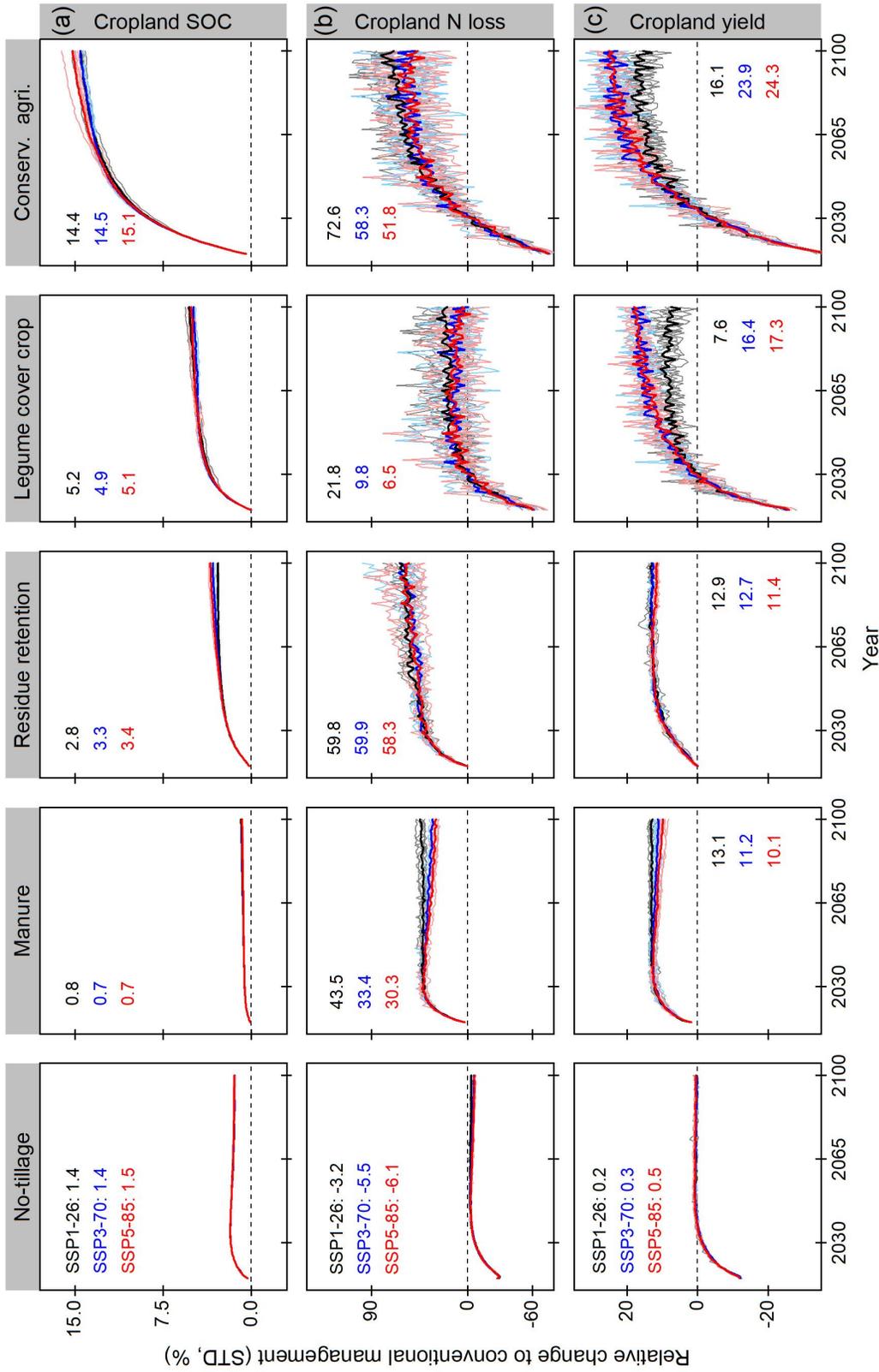
Figure 5.6 The simulated future response (%) of cropland SOC (a), N loss (b) and yield (c) to alternative management practices, relative to the conventional management in eastern Africa (5-C2 runs; Table 2.5). The dark black, blue and red lines denote the mean of simulations using five GCMs for SSP1-26, 3-70 and 5-85 scenarios, respectively. Lines in light color represent the simulation driven by individual GCMs. Numbers in plots indicate the averaged results between 2091 and 2100.
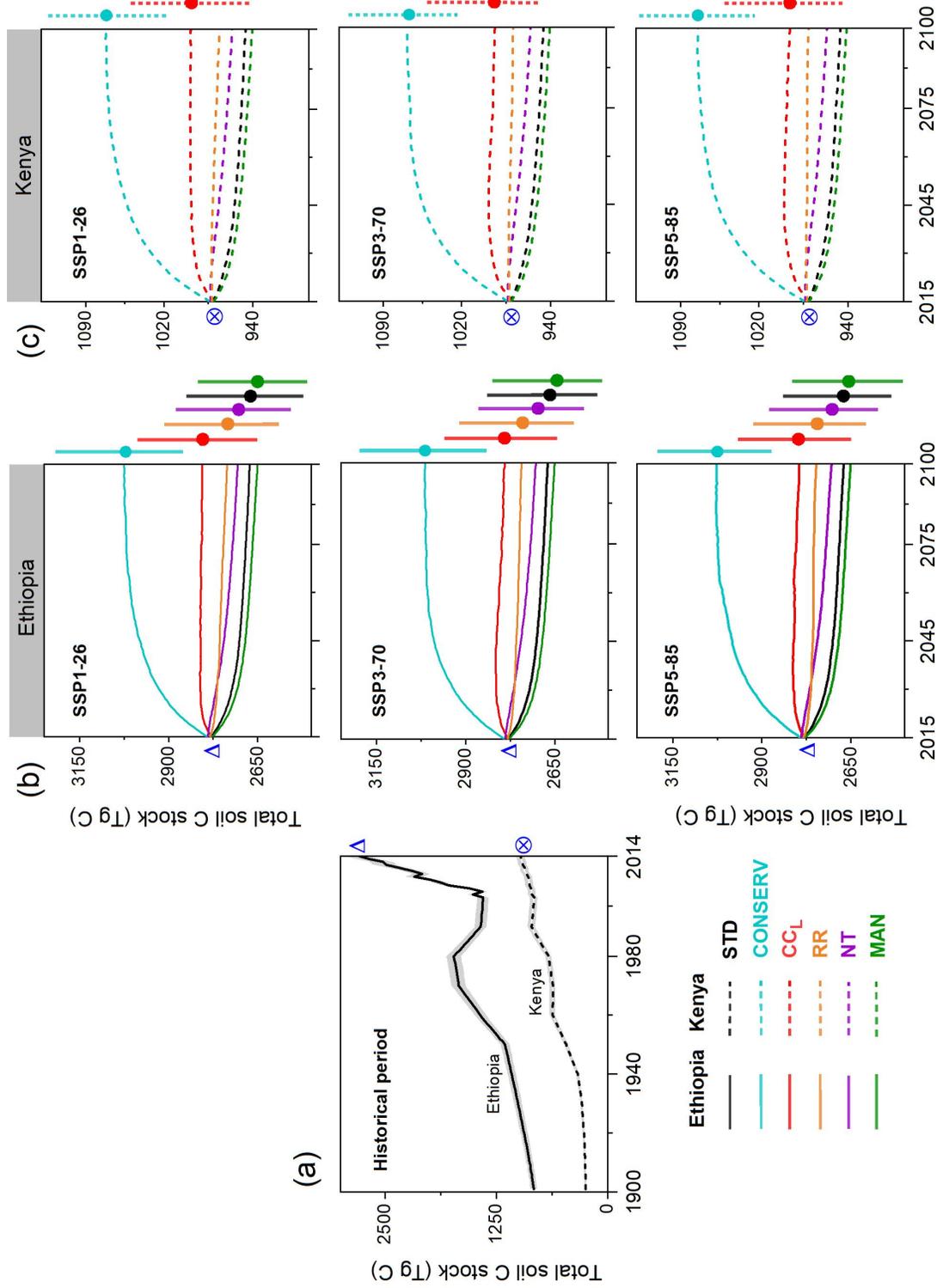
Figure 5.7 The simulated cropland soil C stocks from the conventional management for GCM-based historical period (5-C1 runs; Table 2.5) (a), and SOC potential response to alternative management practices under future SSP1-26, 3-70 and 5-85 scenarios (5-C2 runs; Table 2.5) in Ethiopia (b) and Kenya (c). Blue triangle for Ethiopia and sun cross for Kenya denote the seamless connection between historical period and future simulations. Filled circles represent the averaged result between 2091 and 2100, with a range from five GCMs given in vertical bar. See Fig. 5.4 for management abbreviations.

Given the Fig. 5.7, the simulated cropland SOC stocks (0-150 cm) under future conditions varied widely among the evaluated management practices, with the integrated conservation agriculture as the only strategy that exhibited positive C sequestration over the entire simulation time frame (2015-2100). Practicing N-fixing cover crops was simulated to enhance SOC stocks in the first two decades, after which stable SOC for SSP1-26 and slight C loss for SSP3-70 and 5-85 scenarios were found. Other practices, like the conventional management and no-tillage, showed an obvious declining trend in total C storage from 2015-2100. In addition, there were substantial shifts in the optimal C sequestration practice for the future scenarios, with ~30% of cropland areas in Kenya and Ethiopia (Table 5.2) having the potential transitions at the end of this century in comparison with the present-day climate (5-C2 vs. 5-C3 runs; Table 2.5). Most of these shifts were simulated to come from the other management options to legume cover crop, such as manure application and residue retention (Table 5.2).

5.2 Discussion

5.2.1 Uncertainties on model evaluation at site scale

Simulations from LPJ-GUESS compares well with the average maize yields from observed treatments during the experimental period (2004-2015), but the measured interannual variability of the yields for the evaluated management treatments was not reproduced well. The poor performance in modelling yield variability is likely due to the precipitation discrepancy between the gridded climate input data and field-based weather records (not shown), leading to the effects of extreme weather events being difficult to account for. Also, these impacts of extremes on physiological processes such as flowering or grain filling so far are not well represented in crop models, including LPJ-GUESS, but known to cause yield losses (Olin et al., 2015a; Nyawira et al., 2021).

Multi-cropping within a year has not been implemented to LPJ-GUESS at the moment, absence of maize residue and manure application events in the second cropping season (i.e., the short rainy season from September to January) might thus contribute to underestimating the measured SOC when evaluating the treatments associated with these two practices. In addition, compared with the fixed amount of maize residue left in the field (2 t ha^{-1}), using 50% of residue retention in the model parameterization is not equivalent to these stable and continuous C inputs to soils due to the varying biomass of simulated maize residue between years. This can in part explain the differences in the rates of SOC loss between the observed and simulated values at both sites.

Despite more than 10 years of implementing the improved management practices, the negative soil C sequestration rates were unexpectedly found in both model simulations and filed measurements from 2004-2015. Both INM3 and CT1 sites were under natural grassland prior to the start of the experiments (5-A1 and 5-A2 runs; Table 2.5), hence SOC losses during the experimental period reflected (1) that grassland soils tend to store more carbon than cropland, and (2) that a new SOC equilibrium has not been achieved in the maize systems after 10+ years of management (Lal, 2008). A similar trend was reported by Moebius-Clune et al. (2011), who showed that SOC in western Kenya was still declining even after more than 50 years of conversion from primary forest to maize. In addition, rapid turnover of the SOM in tropics may be another factor influencing SOC trends because of the prevailing warm and moist climate. The turnover-driven C losses at the sites may exceed C increases due to manure and residue application (Kihara et al., 2020; Nyawira et al., 2021).

LPJ-GUESS in general underestimated the rates of SOC loss at the two experimental sites (Figs. 5.1-5.2). Previous studies have shown that high termite activity in western Kenya can significantly accelerate litter decomposition rates in the no-till maize system (Ayuke et al., 2011; Kihara et al., 2015). We do not know if this particular process played an important role at the field trials, but it is not included in the model representation of SOM decay. In principle, decomposition by soil animals could be addressed by adjusting the decomposition parameters in the structural and metabolic litter pools (Nyawira et al., 2021), but adopting such an approach is currently prevented by the lack of assessment information.

To compare the simulated SOC stocks with observations, we scaled up the measured SOC in the upper soil (0-15 cm) to the modelled depth of 150 cm using a simple extrapolation function. However, the extrapolated SOC values are likely to be different from observations at 0-150cm depth due to the varying management effects on SOC changes with depth. For example, a recent analysis suggested that an intermediate and high intensity of tillage can greatly decrease SOC storage in agricultural soils, but large variations existed between soil layers (Haddaway et al., 2017). The scaling of SOC stocks with depth in the analysis cannot reflect this variation, and introduces uncertainties on soil carbon estimates in our assessment.

5.2.2 Crop production at regional scale

Our simulated maize yields at the national level matched FAO statistics in Kenya and Ethiopia, but productions for most other crop types were generally overestimated (Fig. 5.3). One factor contributing to the overestimation is that LPJ-GUESS uses a harvest efficiency of 90% to adjust the

simulated crop yields on large spatial scales (see Sect. 2.1). This value was chosen to account for the crop post-harvest losses caused by mechanical and/or manual damage during harvest operation, or poor handling and/or storage conditions (Stathers et al., 2020). FAO (2011) reports that cereal production losses vary widely across regions due to differences in management techniques, ranging from 5-7% in Europe and North America to 18% in sub-Saharan Africa (SSA). If the reported losses for SSA also apply to Kenya and Ethiopia, the value of 90% implemented in the model would result in a 10% overestimation of production in the region.

Pulse productions in both countries were largely overestimated in the model simulations (Fig. 5.3). This is mainly due to the high N fixation rate in legumes simulated by LPJ-GUESS under warm and humid climate (see evaluation results in Chapter 3). High BNF rate may reduce the N limitations to leaf photosynthesis and subsequently enhance the carbon assimilation flow to storage organ, leading to the high production in N-fixing crops. Similar to pulses, sorghum yields we simulated at the national level were also significantly higher than FAO records (Fig. 5.3). This suggest that other factors are at play as well. For example, insect pests, particularly shoot flies and stalk borers, have been considered to be the major constraint to sorghum production in SSA (Wortmann et al., 2009), reducing yields by an estimate of 11-49% in western Africa and 15-88% in eastern Africa (Okosun et al., 2021). The present LPJ-GUESS crop module does not account for insect pests, which may lead to the large overestimation of sorghum production in our studied region. Additionally, a good representation of photosynthate allocation to various plant organs is particularly important when modelling crop yields (Bondeau et al., 2007). In this chapter we updated the daily assimilate partitioning scheme for sorghum based on the existing literature (see Sect. 2.5.2), but this process has not been parameterized and calibrated against observations from field experiments. Whether this related to the large-scale yield overestimation needs to be further investigated in future work.

5.2.3 Impacts of management practices on crop ecosystems

5.2.3.1 Soil carbon stocks

Statistic-based estimates of management improvement impact on the potential cropland SOC increase in Kenya and Ethiopia range from 15.5-32.7 Tg yr^{-1} assuming that improved managements are continuously implemented over 20 years (Zomer et al., 2017). Across the eastern African study region, LPJ-GUESS produced a SOC enhancement of 2.9 Tg yr^{-1} under the optimal C management (OPT) and 4.7 Tg yr^{-1} under the integrated conservation agriculture practice (CONSERV) compared with the conventional management (STD; Table 5.1). The difference between the estimates in

Zomer et al. (2017) and our simulation may be attributed to our longer model experimental period. When a change in management causes soil C stock to increase, it moves towards a new equilibrium value over a period of years or decades depending on climate and soil type (Johnston et al., 2009; Sommer & Bossio, 2014). In the early years after the change in management the annual rate of increase is largest, and it then gradually declines when the new SOC equilibrium value is achieved (see ΔSOC_{rate} in Fig. 4.4 in Chapter 4). The 86 years of simulations in the model experiment is around four times longer than the 20 years studied in Zomer et al. (2017). If we consider the rates of SOC sequestration over the first 20 years of simulation, the modelled soil C increase of 13.6 Tg yr^{-1} from CONSERV practice (not shown) is close to the lower end of the range reported in Zomer et al. (2017).

Regionally, our simulated small SOC increase of 2% under no-tillage (NT) and 3% under residue retention (RR) agree with a recent meta-analysis of Githongo et al. (2021), in which converting from a conventional tillage to a no-till system in SSA on average showed only slight SOC increase in a maize cropping system. This reported insignificant impact contrasts with an earlier synthesis conducted by Powlson et al. (2016), who reported that the combination of minimum tillage and residue retention in SSA would result in a net SOC increase of 0.45 Mg C ha^{-1} yr^{-1} after three to nine years of implementation, ~24% higher than the control management (i.e., tillage and residue removal). In our model experiments, only the integrated conservation agriculture practice (CONSERV) results in a fairly large SOC increase of 11% (varying from 4-22%; Fig. 5.4), more comparable but still below the findings in Powlson et al. (2016). The reason for the disagreement between the regional simulation and field-based experiments is difficult to assess because of the difference in the studied geographical scales, land-use history, sampled soil depth and implemented duration of practices.

The regional-scale simulated results are consistent with a recent meta-analysis finding that N-fixing cover crops contribute more to increasing SOC storage than non-legume plants (Abdalla et al., 2019). However, it should be noted that in LPJ-GUESS we assumed that cover crops in eastern Africa are rotated with the main crops and thus solely grown during the short rainy season. This assumption is likely to cause cover crop biomass addition to the soil pools being too high since we overestimate the length of the bare-fallow period for cover crop cultivation (Porwollik et al., 2022). Compared to the conventional management with no cover crops, such an overestimation would then be possibly reflected in high SOC sequestration rates. At present more than 90% of total annual

crop yields in Ethiopia are achieved in the long rainy season (Central Statistical Agency, 2016); nevertheless, few farmers are willing to adopt a "main crop + cover crop" rotation pattern as this practice will sacrifice one season of maize production. Our model simulations support earlier findings that implementing legume cover crops solely in the short rainy season is expected to achieve SOC enhancement and may gain yield benefits in the tropics of SSA (Rao & Mathuva, 2000; Carsky et al., 2001), although at some smallholder farms yield increases from N-fixing cover cropping may not completely compensate for the production loss of the short rainy season (Carsky et al., 2001).

The opinion that applying conservation agriculture techniques can enhance cropland SOC stocks is often based on comparing differences among management practices, but without a time perspective (Martinsen et al., 2019; Kihara et al., 2020). The future projections done here show that the measured SOC loss in 12-year experiments in western Kenyan (Fig. 5.2) would continue to take place in other parts of eastern Africa under most evaluated management practices, consistent with the finding of a recent modelling study in Kenya (Nyawira et al., 2021). The 4p1000 initiative launched at COP 21 sets a target of 3.4 Pg C yr^{-1} SOC sequestration in agricultural soils (0-40cm) worldwide to mitigate climate change (Corbeels et al., 2019). However, our modelling results indicate that croplands situated in eastern Africa can potentially achieve this target only if the combined management practice (i.e., CONSERV) would be adopted and sustained. But even if alternative managements may not always support a positive SOC sequestration regionally, they nonetheless are here projected to mitigate soil C losses and show co-benefits for crop production (Kihara et al., 2020).

5.2.3.2 Cropland N loss and yields

Our model simulated a lower N loss of 45-134 Gg N yr^{-1} in Kenya, in comparison with the statistics-based estimates of 76-297 Gg N yr^{-1}, using a nitrogen-budget method (Zhang et al., 2021). Likewise, our conventional management experiment (STD) predicted a regional mean N loss of 9.2 kg N ha^{-1} yr^{-1} (Table 5.1), again below the finding of 16.7-18.2 kg N ha^{-1} yr^{-1} reported in Kaltenegger et al. (2021). One possible explanation for these differences may be missing processes in LPJ-GUESS, such as N loss through surface runoff and soil erosion. Additionally, the nitrogen-budget method used by Zhang et al. (2021) and Kaltenegger et al. (2021) assumed that all crop residues were incorporated to the soils after harvest, almost contrasting with the parameterization in our STD simulation (only 10% of residue retention; Table 2.7). Removing most residues from

cropland in the model experiment is expected to produce low N loss because of less N inputs to the soils compared with 100% of residue retention. To leave all residues in the field, the model regionally produced N loss of 20.8 Gg N yr^{-1} (RR; Table 5.1), comparable with the findings in Kaltenegger et al. (2021).

Compared with the conventional management (STD), an increase of 89% in N losses was simulated over eastern Africa as a result of 100% of residue retention (Fig. 5.4). A global meta-analysis showed that returning residues to the field may enhance gaseous N emissions by 8%-37%, but the same study also pointed to that straw application reduced hydrological N losses by 10%-26% (Xia et al., 2018). At the moment soil hydraulic properties in response to residue management has not been represented in LPJ-GUESS. This missing process is likely to cause the overestimated hydrological N losses since straw return usually mitigates N leaching via enhancing soil water retention in reality (Blanco-Canqui et al., 2007). Additionally, crop residues after harvest in eastern Africa are expected to rapidly decompose as a result of the warm and humid climate, continuously releasing mineralized N for subsequent crop uptake (Kihara et al., 2015). As discussed earlier, only a single growing season within a year was simulated by LPJ-GUESS, the bare-fallow period under the modelled residue management would increase N losses since the reactive N is not used to plant growth in the short rainy season. Nonetheless, simulated crop yields due to residue application still on average increased by 18% regionally (Fig. 5.4) by reason of the enhanced size of the mineral N pool. This result agrees with two previous studies that indicated a yield gain of 19%-35% in Ethiopian wheat systems with 66% of residue return (Adimassu et al., 2019), and another reporting -1%-39% of maize yield changes caused by residue application in semi-arid Kenya (Kihara et al., 2011).

Practicing cover cropping systems enhanced N losses in some simulated grid cells, particularly for legume cover crops (CC$_L$; Fig. 5.4). One possible reason for this simulation is that the enhanced available N derived from the fast decomposition of cover crops would serve as substrate for N losses instead of being taken up by the main crop, as a result of the temporal inconsistency between periods of soil N mineralization and high N demand of the main crop (Marcillo & Miguez, 2017). Compared with the bare-fallow simulation under the conventional management (STD; Table 2.7), LPJ-GUESS produced a small yield gain of 6% under non-legume cover cropping systems but a large increase of 19% in legume cover crops (Fig. 5.4), agreeing with the meta-study findings that using N-fixing cover crops often contributes more to benefiting subsequent crop yields than non-

legumes when N fertilizer investments to main crops are low (Quemada et al., 2013; Marcillo & Miguez, 2017; Thapa et al., 2018). These small simulated yield benefits (even decreased production in few grid cells; see Fig.5.4) from non-legume cover cropping mostly resulted from indirect competition for water and nutrients, which may be unavailable for the following crops cultivated in the long rainy season.

5.3 Conclusions

Using the updated LPJ-GUESS model version from Chapter 3, this chapter presented a large-scale modelling study in eastern Africa, highlighting long-term potential influence of improved management practices on crop ecosystems under various climate change scenarios. The model performance was evaluated against observations from two maize field trials (>10 years implementation of conservation agriculture) in western Kenya and large-scale reported estimates from published sources. Our results showed that crop ecosystems simulated by LPJ-GUESS realistically responded to alternative management strategies and climate variation, and generated the comparable SOC storage, N losses (including gaseous emissions and N leaching) and crop productivity with measurements in the studied region.

Our regional simulations revealed that the improved managements practiced in agricultural soils over eastern Africa can contribute to climate change mitigation while benefiting in crop production, in particular for the integrated conservation agriculture practice with all soil-conserving techniques included. Under the present climate condition, this combined strategy was simulated to increase total soil C storage by 11% in the long term, accompanied by a 25% enhancement in total crop production, in comparison with the conventional management. Adopting legume cover cropping in the model simulations was also identified as a promising individual practice in eastern Africa to increase cropland soil C levels (+4%) and agricultural production (+16%), but with environmental cost of increased total N losses (+28%). These management effects could be also sustained in simulations of three future climate change scenarios. However, it should be noted that absence of processes in the model, such as multi-cropping system and N losses through runoff and soil erosion, might regionally have biased the evaluated management effects on agricultural ecosystems.

6 General conclusions and outlook

This thesis studied the importance and effects of improved management practices—in particular the cultivation of N-fixing legumes—on agricultural ecosystems using the LPJ-GUESS DGVM. The following paragraphs aim to summarize the key findings of this thesis, highlight major sources of uncertainty, and propose possible directions of future model development.

6.1 Answers to the underlying research questions

Chapters 3-5 give answers to the main research questions addressed in this thesis.

– How much nitrogen is fixed from the atmosphere in grain legumes globally?

At a global scale, annual biological N fixation is modelled to be 11.6±2.2 Tg N for soybean and 5.6±1.0 Tg N for all pulses, with a total fixation rate of 17.2±2.9 Tg N yr^{-1} in all grain legumes for the period 1981-2016. The highest BNF rates are found in tropical and temperate regions with warm and humid climates. Soil water availability and temperature are most important factors controlling the N fixation, in addition to N fertilizer application. On a global average, fixing N from the atmosphere is the main source for meeting the N demand of legumes, and contributes 57±4% and 60±1% to total N uptake in soybean and pulses, respectively, for the period 1981-2016.

– Can legume cover crops contribute to environmental sustainability without compromising crop production in global croplands?

Yes, the improved LPJ-GUESS model estimates that the impacts of N-fixing cover crops on global agricultural lands can be beneficial for environmental sustainability without compromising crop production. This potential is highly dependent on location, main-crop type, and management duration. If all current croplands in the world were under conservation agriculture systems, combining legume cover cropping with no-tillage resulted in increased soil carbon levels by 7% (+0.32 Pg C yr^{-1} in global croplands) while reducing N losses by 41% (-7.3 Tg N yr^{-1}) after 36 years of implementation. This integrated practice is accompanied by a potential 2% increase in total crop production (+37 million tonnes per year including wheat, maize, rice, and soybean) in the last decade of the simulation. The identified effect on global SOC sequestration (i.e., 0.32 Pg C yr^{-1} = 1.17 Pg CO_2eq yr^{-1}) would compensate for 22% of annual direct GHG emissions from crop and livestock activities (5.3 Pg CO_2eq yr^{-1} in 2018; FAO, 2020), or is equivalent to 29% of GHG emissions from agricultural land use change (4.0 Pg CO_2eq yr^{-1} in 2018; FAO, 2020).

– **How significant are these impacts from cover crops compared with other agricultural practices under present and future climate conditions?**

Taking eastern Africa as a case study, the six assessed different management practices (which address a range of aspects of sustainable land management) result in a net increase in simulated cropland soil carbon compared with conventional management. Most of the six practices (with the exception of cover crops in some regions) also achieve the extra benefit of increased yields, even through simulations show larger N losses (including gaseous emissions and N leaching). When integrated over Kenya and Ethiopia, the combined conservation agriculture practice—including no-tillage, residue and manure application, and cover cropping—increases total simulated SOC stocks by 11% in the long term, accompanied by a 25% enhancement in total crop production. Practicing legume cover crops in simulations is also identified as a promising individual practice to increase cropland soil C levels (+4%) and agricultural production (+16%), but with environmental cost of increased total N losses (+28%). These management impacts would be also sustained in simulations of three future climate pathways (SSP1-26, 3-70 and 5-85 scenarios taken from five GCMs) over eastern Africa.

6.2 Limitations and future work

6.2.1 Modelling N fixation in legumes

The challenges of modelling N fixation in legumes are mainly due to their large variation in species, sites, and managements. Symbiotic nitrogen fixation by rhizobia is a complex natural process, which is related not only to the N status of host plants and soils in the macro-environment, but also to the process of *Rhizobium* or *Bradyrhizobium* bacteria in root nodules in the micro-environment (Rice et al., 2000). It is hard to incorporate these two different but highly associated processes into one model (Liu et al., 2011; Chen et al., 2016). Furthermore, there is insufficient observed information to establish a reliable relationship between BNF and other factors such as soil pH (Vanlauwe et al., 2019), inoculation effectiveness (Denton et al., 2017), salinity (Zahran, 1999; Bruning & Rozema, 2013), and phosphorus availability (Le Roux et al., 2009; Singh et al., 2012), which are at the moment absent in LPJ-GUESS and other crop models, although many field trials have demonstrated their importance.

Adding mineral N to the soil in LPJ-GUESS can increase soil N uptake, reducing the plant's N deficit and therefore also reducing the upper limit of daily N fixation rate. Although the modelled negative relationship between fertilizer application rates and N fixation showed a generally good

agreement with the observed response across a range of field sites, the simulated BNF rates at the high-fertilized trials were higher than the measured values (Fig. 3.2). This might be partially explained by the underestimation in soil N uptake under high N concentration, resulting in plant N demand remaining very high and substantial N still being fixed. The large discrepancies between modelled and observed N uptake in the high-fertilized treatments suggest that the N uptake representation in LPJ-GUESS should be further improved. A step forward could be to incorporate the inhibitory effects of soil mineral N content on N fixation into the model (Chen et al., 2016; Wu et al., 2020), since experimental evidence indicates that high soil mineral N not only affects plant N uptake in roots, but also depresses legume nodule initiation, nodule size and specific nodule activity, therefore reducing the amount of N fixation from the atmosphere (Herridge et al., 1984; Purcell & Sinclair, 1990; Thornley & Cannell, 2000).

Although high soil-reactive N concentration limits root nodulation and further affects N fixation (Mourtzinis et al., 2018; Brar & Lawley, 2020), moderate soil N levels during the legume vegetative stage favors root growth and nodule formation and stimulates N fixation (Salvagiotti et al., 2008). In field trials, a specific threshold of soil N concentration above (below) which N fixation is inhibited (stimulated) is difficult to measure. Moreover, the timing of N application remains challenging. Studies reported that using N fertilizer as starter N at sowing can increase yields because of sufficient soil reactive N to stimulate soybean growth in the early season (Osborne & Riedell, 2011; Gai et al., 2017). However, other studies argued that early reproductive growth stages are the best time to apply fertilizer, as this is when legumes have the greatest N demand for grain development, soil N reserves are depleting, and N fixation begins to slow (Mourtzinis et al., 2018; Córdova et al., 2019). Unfortunately, there are no consistent conclusions regarding these factors, making it difficult to parameterize the mechanistic processes or setups in LPJ-GUESS at present.

6.2.2 Impacts of cover cropping on crop ecosystems

A detailed evaluation of simulating cover crop (CC) impacts on croplands worldwide remains challenging due to various cover crop species, farming rotation systems, and managements in the field trials. We mainly examined the model performance by categorizing herbaceous CCs as non-legume and legume functional types, with site-specific management practices considered (Fig. 2.3). Although the updated LPJ-GUESS model version from Chapter 3 can reproduce the observed responses of ecosystem service indicators to CC cultivation, the magnitude of these changes did not always match experimental measurements (Fig. 4.1). This likely reflects differences between highly

controlled field conditions and model parameterization in land-use history, initial SOM levels, cropping system managements, and the C-N allocation scheme in CCs. In addition, important processes that determine CC impacts in the field experiments—such as weeds (Mazzoncini et al., 2011), intercropping (Valkama et al., 2015), erosion (Daryanto et al., 2018), and soil structural modification via grass roots (Nouri et al., 2022)—have not been accounted for in the model.

Legume CCs are usually identified as a promising strategy to substitute synthetic N fertilizer in agricultural production due to their high N fixation rates (Peoples et al., 2021; Herridge et al., 2022). Our modelled N fixed by natural C_3 grass (a surrogate for white clover; see Sect. 2.2) during off-season periods is 30-70 kg N ha^{-1} yr^{-1} in warm and humid regions (Fig. 4.9), which is lower than the reported range of 49-154 kg N ha^{-1} yr^{-1} but these latter estimates were for the entire year (Ledgard et al., 2001; Burchill et al., 2014; Anglade et al., 2015). Nonetheless, in our simulations employing legume CCs results in higher yield benefits in the humid tropics compared with non-legumes (Fig. 4.6). As introduced in Sect. 2.1, one main growing season within a year is modelled in LPJ-GUESS, total agricultural production achieved by multi-cropping systems in the tropics are not yet captured. As a consequence, N fixation rate and biomass in legume CCs may be too high since we overestimate the length of the off-season period for cover crop cultivation (Porwollik et al., 2022). Compared with controls with no CCs, such an overestimation would then be possibly reflected in high SOC sequestration rates and yield benefits in tropical climates.

Rather than employing herbaceous CCs, it is more common to use legume crops (e.g., faba bean and field peas) as "green manure" in some temperate regions (Rinnofner et al., 2008; Andersen et al., 2020). These grain legumes are often intercropped with other cash crops, and incorporated to soils at full bloom stage to maximize N fixation rates while minimizing soil water depletion (Williams et al., 2014; Denton et al., 2017). To better represent region-specific cover crop practices, the implementation of N-fixing grain legumes as intercrops, together with multi-cropping systems within a year, remains to be taken into account in future model work.

6.2.3 Trade-offs and win-win strategy in eastern Africa

Taking eastern Africa as a case study, we attempted to identify synergistic management strategies for achieving environmental sustainability and ensuring food security at a large spatial scale. None of the assessed managements accomplished a win-win situation in terms of enhancing soil carbon stocks and crop production while minimizing N losses when integrated over the study regions.

Synergies and trade-offs among the three examined indicators varied between locations and cropping systems.

From the perspectives of food demand and SOC sequestration only, conservation agriculture (CA)—a combined management with no-tillage, residue and manure application, and N-fixing cover crops—is found to be the most promising practice in our simulations for both present-day and future climate conditions. Nonetheless, considering the potential yield loss in the first several years under CA systems (Stevenson et al., 2014; Pittelkow et al., 2015), it may be difficult to convince farmers to employ such a practice in reality, if indeed 1-25% of yield reduction could be expected compared with the conventional management. Farmers would suffer economic losses despite the accompanying 1-10% of increase in the simulated SOC stocks (Fig. 5.6). Furthermore, labor demand and cost-ineffective investment in CA maintenance may prevent this practice from being adopted widely in the study regions (Thierfelder et al., 2013; Kihara et al., 2020). To change this situation, a payment scheme for carbon sequestration legislated by the government or volunteered by corporations and individuals (Salzman et al., 2018) may be needed to compensate for farmer's economic losses, particularly in the context of future climate change.

6.3 Final remarks

Increasing crop productivity while keeping detrimental side-effects on the environment low is a major challenge for global agriculture today. This thesis highlights the possibility of implementing improved agricultural management to mitigate climate change and support food security in global croplands. It also reveals the importance of the cultivation of N-fixing legumes when targeting the long-term sustainable development of agricultural production, particularly in the context of future climate change.

7 Acknowledgments

I have come a long way to get this thesis to you. At the end I would like to thank all those people who made this work possible. In particular, I want to thank the following people:

First of all I am very grateful to my supervisor Almut Arneth who spent countless hours on the success of this thesis despite her numerous other duties. Almut, thank you so much for your guidance, inspiration, patience, and shared knowledge over the last years, particularly for your endless encouragement when I was at the most helpless moment three years ago. I would not have been able to finish this thesis without your support. You have been a great supervisor.

I am also very thankful for my co-advisors Peter Anthoni and Sam Rabin for their help with the LPJ-GUESS model codes, their many suggestions on structuring, writing, and presenting the work.

The same is also true for my co-advisor Anita Bayer. Again, thanks a lot for her help, recommendations, motivation, and the time she invested into collaborated papers.

Special thanks go to Stefan Olin for his help with model development and to Sylvia Nyawirai for shared valuable field-measured data in Kenya for model evaluation.

I also want to thank the other members of the LEMG group at KIT/IMK-IFU, the LPJ-GUESS developers at Lund University, and all of our collaborators within the GIZ-funded project.

Furthermore, I am grateful to Sönke Zaehle, Anja Rammig, Stefan Hinz, and Mark Rounsevell for reviewing this thesis.

Finally, I want to thank my friends and my family for their unlimited company and support, especially in tough times of Covid-19 over the past three years.

References

Abdalla, M., Hastings, A., Cheng, K., Yue, Q., Chadwick, D., Espenberg, M., et al. (2019). A critical review of the impacts of cover crops on nitrogen leaching, net greenhouse gas balance and crop productivity. *Global Change Biology*, *25*(8), 2530–2543. https://doi.org/10.1111/gcb.14644

Abu-shakra, S. S., Phillips, D. A., & Huffaker, R. C. (1978). Nitrogen Fixation and Delayed Leaf Senescence in Soybeans. *Science*, *199*(4332), 973–975.

Adimassu, Z., Alemu, G., & Tamene, L. (2019). Effects of tillage and crop residue management on runoff, soil loss and crop yield in the Humid Highlands of Ethiopia. *Agricultural Systems*, *168*, 11–18. https://doi.org/10.1016/j.agsy.2018.10.007

Andersen, B. J., Samarappuli, D. P., Wick, A., & Berti, M. T. (2020). Faba bean and pea can provide late-fall forage grazing without affecting maize yield the following season. *Agronomy*, *10*(1). https://doi.org/10.3390/agronomy10010080

Anderson, R. G., Canadell, J. G., Randerson, J. T., Jackson, R. B., Hungate, B. A., Baldocchi, D. D., et al. (2011). Biophysical considerations in forestry for climate protection. *Frontiers in Ecology and the Environment*, *9*(3), 174–182. https://doi.org/10.1890/090179

Anglade, J., Billen, G., & Garnier, J. (2015). Relationships for estimating N_2 fixation in legumes: Incidence for N balance of legume-based cropping systems in europe. *Ecosphere*, *6*(3), 1–24. https://doi.org/10.1890/ES14-00353.1

Arneth, A., Shin, Y. J., Leadley, P., Rondinini, C., Bukvareva, E., Kolb, M., et al. (2020). Post-2020 biodiversity targets need to embrace climate change. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(49), 30882–30891. https://doi.org/10.1073/pnas.2009584117

Arneth, A., Olsson, L., Cowie, A., Erb, K. H., Hurlbert, M., Kurz, W. A., et al. (2021). Restoring Degraded Lands. *Annual Review of Environment and Resources*, *46*, 569–599. https://doi.org/10.1146/annurev-environ-012320-054809

Aronsson, H., Stenberg, M., & Ulén, B. (2011). Leaching of N, P and glyphosate from two soils after herbicide treatment and incorporation of a ryegrass catch crop. *Soil Use and Management*, *27*(1), 54–68. https://doi.org/10.1111/j.1475-2743.2010.00311.x

Arora, V. K., & Montenegro, A. (2011). Small temperature benefits provided by realistic afforestation efforts. *Nature Geoscience*, *4*(8), 514–518. https://doi.org/10.1038/ngeo1182

Austin, K. G., Schwantes, A., Gu, Y., & Kasibhatla, P. S. (2019). What causes deforestation in Indonesia? *Environmental Research Letters*, *14*(2). https://doi.org/10.1088/1748-9326/aaf6db

Ayuke, F. O., Pulleman, M. M., Vanlauwe, B., de Goede, R. G. M., Six, J., Csuzdi, C., & Brussaard, L. (2011). Agricultural management affects earthworm and termite diversity across humid to

semi-arid tropical zones. *Agriculture, Ecosystems and Environment*, *140*(1–2), 148–154. https://doi.org/10.1016/j.agee.2010.11.021

Beltran-Peña, A., Rosa, L., & D'Odorico, P. (2020). Global food self-sufficiency in the 21st century under sustainable intensification of agriculture. *Environmental Research Letters*, *15*(9). https://doi.org/10.1088/1748-9326/ab9388

Blanchy, G., Bragato, G., Bene, C. Di, Jarvis, N., Larsbo, M., Meurer, K., & Garré, S. (2023). Soil and crop management practices and the water regulation functions of soils : a synthesis of meta-analyses relevant to European agriculture. *Soil*, *9*, 1–20. https://doi.org/10.5194/soil-9-1-2023

Blanco-Canqui, H., Lal, R., Post, W. M., Izaurralde, R. C., & Shipitalo, M. J. (2007). Soil hydraulic properties influenced by corn stover removal from no-till corn in Ohio. *Soil and Tillage Research*, *92*(1–2), 144–155. https://doi.org/10.1016/j.still.2006.02.002

Blanco-Canqui, H., Shaver, T. M., Lindquist, J. L., Shapiro, C. A., Elmore, R. W., Francis, C. A., & Hergert, G. W. (2015). Cover Crops and Ecosystem Services: Insights from Studies in Temperate Soils. *Agronomy Journal*, *107*(6), 2449–2474. https://doi.org/10.2134/agronj15.0086

Von Bloh, W., Schaphoff, S., Müller, C., Rolinski, S., Waha, K., & Zaehle, S. (2018). Implementing the nitrogen cycle into the dynamic global vegetation, hydrology, and crop growth model LPJmL (version 5.0). *Geoscientific Model Development*, *11*(7), 2789–2812. https://doi.org/10.5194/gmd-11-2789-2018

Bolinder, M. A., Crotty, F., Elsen, A., Frac, M., Kismányoky, T., Lipiec, J., et al. (2020). The effect of crop residues, cover crops, manures and nitrogen fertilization on soil organic carbon changes in agroecosystems: a synthesis of reviews. *Mitigation and Adaptation Strategies for Global Change*, *25*(6), 929–952. https://doi.org/10.1007/s11027-020-09916-3

Bondeau, A., Smith, P. C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., et al. (2007). Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biology*, *13*(3), 679–706. https://doi.org/10.1111/j.1365-2486.2006.01305.x

Boote, K., Mínguez, M., & Sau, F. (2002). Adapting the CROPGRO legume model to simulate growth of faba bean. *Agronomy Journal*, *94*(4), 743–756. https://doi.org/10.2134/agronj2002.7430

Boote, K., Hoogenboom, G., Ingram, K., & Jones, J. (2009). Modeling Nitrogen Fixation and Its Relationship to Nitrogen Uptake in the CROPGRO Model. In L. Ma, L. R. Ahuja, & T. Bruulsema (Eds.), *Quantifying and Understanding Plant Nitrogen Uptake for Systems Modeling* (pp. 13–46). https://doi.org/10.1201/9781420052978

Brar, N., & Lawley, Y. (2020). Short-season soybean yield and protein unresponsive to starter nitrogen fertilizer. *Agronomy Journal*, *112*(6), 5012–5023. https://doi.org/10.1002/agj2.20378

Bruning, B., & Rozema, J. (2013). Symbiotic nitrogen fixation in legumes: Perspectives for saline agriculture. *Environmental and Experimental Botany*, *92*, 134–143. https://doi.org/10.1016/j.envexpbot.2012.09.001

Burchill, W., James, E. K., Li, D., Lanigan, G. J., Williams, M., Iannetta, P. P. M., & Humphreys, J. (2014). Comparisons of biological nitrogen fixation in association with white clover (Trifolium repens L.) under four fertiliser nitrogen inputs as measured using two 15N techniques. *Plant and Soil*, *385*(1–2), 287–302. https://doi.org/10.1007/s11104-014-2199-1

Cabelguenne, M., Debaeke, P., & Bouniols, A. (1999). EPICphase, a version of the EPIC model simulating the effects of water and nitrogen stress on biomass and yield, taking account of developmental stages: Validation on maize, sunflower, sorghum, soybean and winter wheat. *Agricultural Systems*, *60*(3), 175–196. https://doi.org/10.1016/S0308-521X(99)00027-X

Camargo-Alvarez, H., Elliott, R. J. R., Olin, S., Wang, X., Wang, C., Ray, D. K., & Pugh, T. A. M. (2022). Modelling crop yield and harvest index: the role of carbon assimilation and allocation parameters. *Modeling Earth Systems and Environment*. https://doi.org/10.1007/s40808-022-01625-x

Canadell, J. G., Monteiro, P. M. S., Costa, M. H., Cotrim da Cunha, L., Cox, P. M., Eliseev, A. V., et al. (2021). Global Carbon and Other Biogeochemical Cycles and Feedbacks. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 673–816). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. https://doi.org/10.1017/9781009157896.007

Carsky, R. J., Becker, M., & Hauser, S. (2001). Mucuna Cover Crop Fallow Systems: Potential and limitations. In G. Tian, F. Ishida, & D. Keatinge (Eds.), *Sustaining soil fertility in West Africa* (pp. 111–135). SSSA Special Publication. https://doi.org/10.2136/sssaspecpub58

Central Statistical Agency. (2016). The federal democratic republic of Ethiopia central statistical agency, agricultural sample survery 2015/16 (2008 E.C), volume V: report on area, production and farm management practice of belg season crops for private peasant holdings. Statistical Bulletin. Addis Ababa, Ethiopia.

Chatskikh, D., Hansen, S., Olesen, J. E., & Petersen, B. M. (2009). A simplified modelling approach for quantifying tillage effects on soil carbon stocks. *European Journal of Soil Science*, *60*(6), 924–934. https://doi.org/10.1111/j.1365-2389.2009.01185.x

Chen, C., Lawes, R., Fletcher, A., Oliver, Y., Robertson, M., Bell, M., & Wang, E. (2016). How well can APSIM simulate nitrogen uptake and nitrogen fixation of legume crops? *Field Crops Research*, *187*, 35–48. https://doi.org/10.1016/j.fcr.2015.12.007

Ciais, P., Gervois, S., Vuichard, N., Piao, S. L., & Viovy, N. (2011). Effects of land use change and management on the European cropland carbon balance. *Global Change Biology*, *17*(1), 320–338. https://doi.org/10.1111/j.1365-2486.2010.02341.x

Ciampitti, I. A., & Salvagiotti, F. (2018). New insights into soybean biological nitrogen fixation. *Agronomy Journal*, *110*(4), 1185–1196. https://doi.org/10.2134/agronj2017.06.0348

Cleveland, C. C., Townsend, A. R., Schimel, D. S., Fisher, H., Howarth, R. W., Hedin, L. O., et al. (1999). Global patterns of terrestrial biological nitrogen (N_2) fixation in natural ecosystems. *Global Biogeochemical Cycles*, *13*(2), 623–645. https://doi.org/10.1002/(ISSN)1944-9224

Cleveland, C. C., Houlton, B. Z., Smith, W. K., Marklein, A. R., Reed, S. C., Parton, W., et al. (2013). Patterns of new versus recycled primary production in the terrestrial biosphere. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(31), 12733–12737. https://doi.org/10.1073/pnas.1302768110

Constantin, J., Mary, B., Laurent, F., Aubrion, G., Fontaine, A., Kerveillant, P., & Beaudoin, N. (2010). Effects of catch crops, no till and reduced nitrogen fertilization on nitrogen leaching and balance in three long-term experiments. *Agriculture, Ecosystems and Environment*, *135*(4), 268–278. https://doi.org/10.1016/j.agee.2009.10.005

Corbeels, M., Graaff, J. De, Hycenth, T., Penot, E., Baudron, F., Naudin, K., et al. (2014). Understanding the impact and adoption of conservation agriculture in Africa: A multi-scale analysis. *Agriculture, Ecosystems and Environment*, *187*, 155–170.

Corbeels, M., Cardinael, R., Naudin, K., Guibert, H., & Torquebiau, E. (2019). The 4 per 1000 goal and soil carbon storage under agroforestry and conservation agriculture systems in sub-Saharan Africa. *Soil and Tillage Research*, *188*, 16–26. https://doi.org/10.1016/j.still.2018.02.015

Córdova, S. C., Castellano, M. J., Dietzel, R., Licht, M. A., Togliatti, K., Martinez-Feria, R., & Archontoulis, S. V. (2019). Soybean nitrogen fixation dynamics in Iowa, USA. *Field Crops Research*, *236*, 165–176. https://doi.org/10.1016/j.fcr.2019.03.018

Córdova, S. C., Archontoulis, S. V., & Licht, M. A. (2020). Soybean profitability and yield component response to nitrogen fertilizer in Iowa. *Agrosystems, Geosciences & Environment*, *3*(1), 1–16. https://doi.org/10.1002/agg2.20092

Corre-Hellou, G., Faure, M., Launay, M., Brisson, N., & Crozat, Y. (2009). Adaptation of the STICS intercrop model to simulate crop growth and N accumulation in pea-barley intercrops. *Field Crops Research*, *113*, 72–81. https://doi.org/10.1016/j.fcr.2009.04.007

Cucchi, M., P. Weedon, G., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., et al. (2020). WFDE5: Bias-adjusted ERA5 reanalysis data for impact studies. *Earth System Science Data*, *12*(3), 2097–2120. https://doi.org/10.5194/essd-12-2097-2020

D'Odorico, P., Davis, K. F., Rosa, L., Carr, J. A., Chiarelli, D., Dell'Angelo, J., et al. (2018). The Global Food-Energy-Water Nexus. *Reviews of Geophysics*, *56*(3), 456–531. https://doi.org/10.1029/2017RG000591

Dabney, S. M., Delgado, J. A., & Reeves, D. W. (2001). Using winter cover crops to improve soil and water quality. *Communications in Soil Science and Plant Analysis*, *32*(7–8), 1221–1250. https://doi.org/10.1081/CSS-100104110

Daryanto, S., Fu, B., Wang, L., Jacinthe, P. A., & Zhao, W. (2018). Quantitative synthesis on the ecosystem services of cover crops. *Earth-Science Reviews*, *185*, 357–373. https://doi.org/10.1016/j.earscirev.2018.06.013

Davies-Barnard, T., & Friedlingstein, P. (2020). The Global Distribution of Biological Nitrogen Fixation in Terrestrial Natural Ecosystems. *Global Biogeochemical Cycles*, *34*(3), 1–17. https://doi.org/10.1029/2019GB006387

Davies-Barnard, T., Zaehle, S., & Friedlingstein, P. (2022). Assessment of the impacts of biological nitrogen fixation structural uncertainty in CMIP6 earth system models. *Biogeosciences*, *19*(14), 3491–3503. https://doi.org/10.5194/bg-19-3491-2022

DeAngelis, A., Dominguez, F., Fan, Y., Robock, A., Kustu, M. D., & Robinson, D. (2010). Evidence of enhanced precipitation due to irrigation over the Great Plains of the United States. *Journal of Geophysical Research*, *115*, 1–14. https://doi.org/10.1029/2010JD013892

Denton, M. D., Phillips, L. A., Peoples, M. B., Pearce, D. J., Swan, A. D., Mele, P. M., & Brockwell, J. (2017). Legume inoculant application methods: effects on nodulation patterns, nitrogen fixation, crop growth and yield in narrow-leaf lupin and faba bean. *Plant and Soil*, *419*(1–2), 25–39. https://doi.org/10.1007/s11104-017-3317-7

Don, A., Schumacher, J., & Freibauer, A. (2011). Impact of tropical land-use change on soil organic carbon stocks - a meta-analysis. *Global Change Biology*, *17*(4), 1658–1670. https://doi.org/10.1111/j.1365-2486.2010.02336.x

Drewniak, B., Song, J., Prell, J., Kotamarthi, V. R., & Jacob, R. (2013). Modeling agriculture in the Community Land Model. *Geoscientific Model Development*, *6*(2), 495–515. https://doi.org/10.5194/gmd-6-495-2013

Duval, M. E., Galantini, J. A., Capurro, J. E., & Martinez, J. M. (2016). Winter cover crops in soybean monoculture: Effects on soil organic carbon and its fractions. *Soil and Tillage Research*, *161*, 95–105. https://doi.org/10.1016/j.still.2016.04.006

Eckersten, H., Geijersstam, L., & Torssell, B. (2006). Modelling nitrogen fixation of pea (Pisum sativum L.). *Acta Agriculturae Scandinavica Section B: Soil and Plant Science*, *56*(2), 129–137. https://doi.org/10.1080/0906471051003050

Elliott, J., Müller, C., Deryng, D., Chryssanthacopoulos, J., Boote, K. J., Büchner, M., et al. (2015). The Global Gridded Crop Model Intercomparison: Data and modeling protocols for Phase 1

(v1.0). *Geoscientific Model Development*, *8*(2), 261–277. https://doi.org/10.5194/gmd-8-261-2015

Erenstein, O. (2003). Smallholder conservation farming in the tropics and sub-tropics: A guide to the development and dissemination of mulching with crop residues and cover crops. *Agriculture, Ecosystems and Environment*, *100*(1–3), 17–37. https://doi.org/10.1016/S0167-8809(03)00150-6

Erisman, J. W., Sutton, M. A., Galloway, J., Klimont, Z., & Winiwarter, W. (2008). How a century of ammonia synthesis changed the world. *Nature Geoscience*, *1*, 636–639. https://doi.org/10.1038/ngeo325

Erisman, J. W., Galloway, J. N., Seitzinger, S., Bleeker, A., Dise, N. B., Roxana Petrescu, A. M., et al. (2013). Consequences of human modification of the global nitrogen cycle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1621). https://doi.org/10.1098/rstb.2013.0116

Etemadi, F., Hashemi, M., Zandvakili, O. R., & Mangan, F. X. (2018). Phenology, Yield and Growth Pattern of Faba Bean Varieties. *International Journal of Plant Production*, *12*(4), 243–250. https://doi.org/10.1007/s42106-018-0023-1

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Fageria, N. K. (2007). Green manuring in crop production. *Journal of Plant Nutrition*, *30*(5), 691–719. https://doi.org/10.1080/01904160701289529

FAO. (2011). *Food loss and food waste: Extent, Causes, and Prevention*. Rome, Italy. Retrieved from https://www.fao.org/3/i2697e/i2697e.pdf

FAO. (2020). *Emissions due to agriculture. Global, regional and country trends 2000-2018. FAOSTAT Analytical Brief Series No 18*. Rome, Italy. Retrieved from https://www.fao.org/3/cb3808en/cb3808en.pdf

FAO, & ITPS. (2015). *Status of the World's Soil Resources– Main Report*. Rome, Italy. Retrieved from https://www.fao.org/3/i5199e/i5199e.pdf

FAOSTAT. (2023). Production/Crops and livestock products [Data set]. Retrieved from https://www.fao.org/faostat/en/#data

Fontaine, D., Kristensen, R. K., Rasmussen, J., & Eriksen, J. (2022). Data on growth, uptake and N_2 fixation of grass-clover leys fertilized with mineral N fertilizer and cattle slurry. *Data in Brief*, *42*, 107998. https://doi.org/10.1016/j.dib.2022.107998

Friedlingstein, P., Jones, M. W., Sullivan, M. O., Andrew, R. M., Bakker, D. C. E., Hauck, J., et al. (2022). Global Carbon Budget 2021. *Earth System Science Data*, *14*, 1917–2005. https://doi.org/10.5194/essd-14-1917-2022

Gai, Z., Zhang, J., & Li, C. (2017). Effects of starter nitrogen fertilizer on soybean root activity, leaf photosynthesis and grain yield. *PLoS ONE*, *12*(4), 1–15. https://doi.org/10.1371/journal.pone.0174841

Garba, I. I., Bell, L. W., & Williams, A. (2022). Cover crop legacy impacts on soil water and nitrogen dynamics, and on subsequent crop yields in drylands: a meta-analysis. *Agronomy for Sustainable Development*, *42*(3). https://doi.org/10.1007/s13593-022-00760-0

Gelfand, I., & Robertson, G. P. (2015). A reassessment of the contribution of soybean biological nitrogen fixation to reactive N in the environment. *Biogeochemistry*, *123*(1–2), 175–184. https://doi.org/10.1007/s10533-014-0061-4

Gichangi, E. M., Karanja, N. K., & Wood, C. W. (2006). Composting cattle manure from zero grazing system with agro-organic wastes to minimise nitrogen losses in smallholder farms in kenya. *Tropical and Subtropical Agroecosystems*, *6*(1870–0462), 57–64.

Githongo, M. W., Kiboi, M. N., Ngetich, F. K., Musafiri, C. M., Muriuki, A., & Fliessbach, A. (2021). The effect of minimum tillage and animal manure on maize yields and soil organic carbon in sub-Saharan Africa. *Environmental Challenges*, *5*, 100340. https://doi.org/10.1016/j.envc.2021.100340

Global Soil Data Task. (2014). Global Soil Data Products CD-ROM Contents (IGBP-DIS). https://doi.org/10.3334/ORNLDAAC/565

Gordon, L. J., Steffen, W., Jönsson, B. F., Folke, C., Falkenmark, M., & Johannessen, Å. (2005). Human modification of global water vapor flows from the land surface. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(21), 7612–7617. https://doi.org/10.1073/pnas.0500208102

Green, J. M. H., Croft, S. A., Durán, A. P., Balmford, A. P., Burgess, N. D., Fick, S., et al. (2019). Linking global drivers of agricultural trade to on-the-ground impacts on biodiversity. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(46), 23202–23208. https://doi.org/10.1073/pnas.1905618116

Gross, A., & Glaser, B. (2021). Meta-analysis on how manure application changes soil organic carbon storage. *Scientific Reports*, *11*(1), 1–13. https://doi.org/10.1038/s41598-021-82739-7

Gu, B., Zhang, X., Lam, S. K., Yu, Y., Grinsven, H. J. M., Zhang, S., et al. (2023). Cost-effective mitigation of nitrogen pollution from global croplands. *Nature*, *613*, 77–84. https://doi.org/10.1038/s41586-022-05481-8

Haddaway, N. R., Hedlund, K., Jackson, L. E., Kätterer, T., Lugato, E., Thomsen, I. K., et al. (2017). How does tillage intensity affect soil organic carbon? A systematic review. *Environmental Evidence*, *6*(30). https://doi.org/10.1186/s13750-017-0108-9

Halliday, J., & Pate, J. S. (1976). The acetylene reduction assay as a means of studying nitrogen fixation in white clover under sward and laboratory conditions. *Grass and Forage Science*, *31*, 29–35. https://doi.org/10.1111/j.1365-2494.1976.tb01112.x

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, *7*, 109. https://doi.org/10.1038/s41597-020-0453-3

Heilmayr, R., Rausch, L. L., Munger, J., & Gibbs, H. K. (2020). Brazil's Amazon Soy Moratorium reduced deforestation. *Nature Food*, *1*(12), 801–810. https://doi.org/10.1038/s43016-020-00194-5

Herridge, D. F., Roughley, R. J., & Brockwell, J. (1984). Effect of rhizobia and soil nitrate on the establishment and functioning of the soybean symbiosis in the field. *Australian Journal of Agricultural Research*, *35*(2), 149–161. https://doi.org/10.1071/AR9840149

Herridge, D. F., Peoples, M. B., & Boddey, R. M. (2008). Global inputs of biological nitrogen fixation in agricultural systems. *Plant and Soil*, *311*(1–2), 1–18. https://doi.org/10.1007/s11104-008-9668-3

Herridge, D. F., Giller, K. E., Jensen, E. S., & Peoples, M. B. (2022). Quantifying country-to-global scale nitrogen fixation for grain legumes II. Coefficients, templates and estimates for soybean, groundnut and pulses. *Plant and Soil*, *474*, 1–15. https://doi.org/10.1007/s11104-021-05166-7

Herzfeld, T., Heinke, J., Rolinski, S., & Müller, C. (2021). Soil organic carbon dynamics from agricultural management practices under climate change. *Earth System Dynamics*, *12*(4), 1037–1055. https://doi.org/10.5194/esd-12-1037-2021

Hu, A., Tang, T., & Liu, Q. (2018). Nitrogen use efficiency in different rice-based rotations in southern China. *Nutrient Cycling in Agroecosystems*, *112*(1), 75–86. https://doi.org/10.1007/s10705-018-9930-x

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., et al. (2020). Harmonization of Global Land Use Change and Management for the Period 2015-2300 (LUH2) for CMIP6. *Geoscientific Model Development*, *13*, 5425–5464. https://doi.org/10.5194/gmd-13-5425-2020

Irmak, S., Odhiambo, L. O., Specht, J. E., & Djaman, K. (2013). Hourly and daily single and basal evapotranspiration crop coefficients as a function of growing degree days, days after emergence, leaf area index, fractional green canopy cover, and plant phenology for soybean. *Transactions of the ASABE*, *56*(5), 1785–1803. https://doi.org/10.13031/trans.56.10219

Van Ittersum, M. K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., & Hochman, Z. (2013). Yield gap analysis with local to global relevance-A review. *Field Crops Research*, *143*, 4–17. https://doi.org/10.1016/j.fcr.2012.09.009

Jägermeyr, J., Gerten, D., Heinke, J., Schaphoff, S., Kummu, M., & Lucht, W. (2015). Water savings potentials of irrigation systems: Global simulation of processes and linkages. *Hydrology and Earth System Sciences*, *19*(7), 3073–3091. https://doi.org/10.5194/hess-19-3073-2015

Jang, W. S., Neff, J. C., Im, Y., Doro, L., & Herrick, J. E. (2021). The Hidden Costs of Land Degradation in US Maize Agriculture. *Earth's Future*, *9*, e2020EF001641. https://doi.org/10.1029/2020EF001641

Jian, J., Du, X., Reiter, M. S., & Stewart, R. D. (2020). A meta-analysis of global cropland soil carbon changes due to cover cropping. *Soil Biology and Biochemistry*, *143*, 107735. https://doi.org/10.1016/j.soilbio.2020.107735

Jobbágy, E. G., & Jackson, R. B. (2000). The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecological Applications*, *10*(2), 423–436. https://doi.org/10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2

Johnston, A. E., Poulton, P. R., & Coleman, K. (2009). Chapter 1 Soil Organic Matter. Its Importance in Sustainable Agriculture and Carbon Dioxide Fluxes. *Advances in Agronomy* (1st ed., Vol. 101). Elsevier Inc. https://doi.org/10.1016/S0065-2113(08)00801-8

Kaltenegger, K., Erb, K. H., Matej, S., & Winiwarter, W. (2021). Gridded soil surface nitrogen surplus on grazing and agricultural land: Impact of land use maps. *Environmental Research Communications*, *3*(5). https://doi.org/10.1088/2515-7620/abedd8

Kaschuk, G., Kuyper, T. W., Leffelaar, P. A., Hungria, M., & Giller, K. E. (2009). Are the rates of photosynthesis stimulated by the carbon sink strength of rhizobial and arbuscular mycorrhizal symbioses? *Soil Biology and Biochemistry*, *41*(6), 1233–1244. https://doi.org/10.1016/j.soilbio.2009.03.005

Kaschuk, G., Hungria, M., Leffelaar, P. A., Giller, K. E., & Kuyper, T. W. (2010). Differences in photosynthetic behaviour and leaf senescence of soybean (Glycine max [L.] Merrill) dependent on N_2 fixation or nitrate supply. *Plant Biology*, *12*(1), 60–69. https://doi.org/10.1111/j.1438-8677.2009.00211.x

Kaspar, T. C., Jaynes, D. B., Parkin, T. B., Moorman, T. B., & Singer, J. W. (2012). Effectiveness of oat and rye cover crops in reducing nitrate losses in drainage water. *Agricultural Water Management*, *110*(3), 25–33. https://doi.org/10.1016/j.agwat.2012.03.010

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., et al. (2020). TRY plant trait database – enhanced coverage and open access. *Global Change Biology*, *26*(1), 119–188. https://doi.org/10.1111/gcb.14904

Kaye, J. P., & Quemada, M. (2017). Using cover crops to mitigate and adapt to climate change. A review. *Agronomy for Sustainable Development*, *37*(1). https://doi.org/10.1007/s13593-016-0410-x

Khan, D. F., Peoples, M. B., Schwenke, G. D., Felton, W. L., Chen, D., & Herridge, D. F. (2003). Effects of below-ground nitrogen on N balances of field-grown fababean, chickpea, and barley. *Australian Journal of Agricultural Research*, *54*(4), 333–340. https://doi.org/10.1071/AR02105

Kihara, J., Martius, C., Bationo, A., & Vlek, P. L. G. (2011). Effects of tillage and crop residue application on Soybean nitrogen fixation in a tropical ferralsol. *Agriculture (Switzerland)*, *1*(1), 22–37. https://doi.org/10.3390/agriculture1010022

Kihara, J., Martius, C., & Bationo, A. (2015). Crop residue disappearance and macrofauna activity in sub-humid western Kenya. *Nutrient Cycling in Agroecosystems*, *102*(1), 101–111. https://doi.org/10.1007/s10705-014-9649-2

Kihara, J., Bolo, P., Kinyua, M., Nyawira, S. S., & Sommer, R. (2020). Soil health and ecosystem services: Lessons from sub-Sahara Africa (SSA). *Geoderma*, *370*(July 2019), 114342. https://doi.org/10.1016/j.geoderma.2020.114342

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan*, *93*(1), 5–48. https://doi.org/10.2151/jmsj.2015-001

Kuo, S., Sainju, U. M., & Jellum, E. J. (1997). Winter Cover Crop Effects on Soil Organic Carbon and Carbohydrate in Soil. *Soil Science Society of America Journal*, *61*(1), 145–152. https://doi.org/10.2136/sssaj1997.03615995006100010022x

Laekemariam, F., Kibret, K., Mamo, T., Karltun, E., & Gebrekidan, H. (2016). Physiographic characteristics of agricultural lands and farmers' soil fertility management practices in Wolaita zone, Southern Ethiopia. *Environmental Systems Research*, *5*(1). https://doi.org/10.1186/s40068-016-0076-z

Lal, R. (2004). Soil carbon sequestration impacts on global climate change and food security. *Science*, *304*(5677), 1623–1627. https://doi.org/10.1126/science.1097396

Lal, R. (2005). World crop residues production and implications of its use as a biofuel. *Environment International*, *31*(4), 575–584. https://doi.org/10.1016/j.envint.2004.09.005

Lal, R. (2008). Soil carbon stocks under present and future climate with specific reference to European ecoregions. *Nutrient Cycling in Agroecosystems*, *81*(2), 113–127. https://doi.org/10.1007/s10705-007-9147-x

Lal, R. (2015). Restoring soil quality to mitigate soil degradation. *Sustainability*, *7*(5), 5875–5895. https://doi.org/10.3390/su7055875

Lamb, D., Erskine, P. D., & Parrotta, J. A. (2005). Restoration of Degraded Tropical Forest Landscapes. *Science*, *310*, 1628–1632. https://doi.org/10.1126/science.1111773

Lan, X., Tans, P., & Thoning, K. W. (2023). Trends in globally-averaged CO_2 determined from NOAA Global Monitoring Laboratory measurements. Version 2023-01 NOAA/GML. Retrieved from gml.noaa.gov/ccgg/trends/

Lange, S. (2019). Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development*, *12*, 3055–3070. https://doi.org/10.5194/gmd-2019-36

Ledgard, S. F., Sprosen, M. S., Penno, J. W., & Rajendram, G. S. (2001). Nitrogen fixation by white clover in pastures grazed by dairy cows: Temporal variation and effects of nitrogen fertilization. *Plant and Soil*, *229*(2), 177–187. https://doi.org/10.1023/A:1004833804002

Lemma, B., Williams, S., & Paustian, K. (2021). Long term soil carbon sequestration potential of smallholder croplands in southern Ethiopia with DAYCENT model. *Journal of Environmental Management*, *294*, 112893. https://doi.org/10.1016/j.jenvman.2021.112893

Lemola, R., & Turtola, E. (2000). Undersowing Italian ryegrass diminishes nitrogen leaching from spring barley. *Agricultural and Food Science in Finland*, *9*(3), 201–215. https://doi.org/10.23986/afsci.5661

Levis, S., Gordon, B. B., Erik Kluzek, Thornton, P. E., Jones, A., Sacks, W. J., & Kucharik, C. J. (2012). Interactive Crop Management in the Community Earth System Model (CESM1): Seasonal influences on land-atmosphere fluxes. *Journal of Climate*, *25*(14), 4839–4859. https://doi.org/10.1175/JCLI-D-11-00446.1

Li, Yan, Zhao, M., Motesharrei, S., Mu, Q., Kalnay, E., & Li, S. (2015). Local cooling and warming effects of forests based on satellite observations. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms7603

Li, Yue, Piao, S., Chen, A., Ciais, P., & Li, L. Z. X. (2020). Local and teleconnected temperature effects of afforestation and vegetation greening in China. *National Science Review*, *7*(5), 897–912. https://doi.org/10.1093/nsr/nwz132

Lin, B. Le, Sakoda, A., Shibasaki, R., & Suzuki, M. (2001). A modelling approach to global nitrate leaching caused by anthropogenic fertilisation. *Water Research*, *35*(8), 1961–1968. https://doi.org/10.1016/S0043-1354(00)00484-X

Lindeskog, M., Arneth, A., Bondeau, A., Waha, K., Seaquist, J., Olin, S., & Smith, B. (2013). Implications of accounting for land use in simulations of ecosystem carbon cycling in Africa. *Earth System Dynamics*, *4*(2), 385–407. https://doi.org/10.5194/esd-4-385-2013

Liu, J., You, L., Amini, M., Obersteiner, M., Herrero, M., Zehnder, A. J. B., & Yang, H. (2010). A high-resolution assessment on global nitrogen flows in cropland. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(17), 8035–8040. https://doi.org/10.1073/pnas.0913658107

Liu, L., Xu, W., Lu, X., Zhong, B., & Guo, Y. (2022). Exploring global changes in agricultural ammonia emissions and their contribution to nitrogen deposition since 1980. *Proceedings of the National Academy of Sciences of the United States of America*, 1–9. https://doi.org/10.1073/pnas.2121998119

Liu, Q., Liu, B., Zhang, Y., Hu, T., Lin, Z., Liu, G., et al. (2019). Biochar application as a tool to decrease soil nitrogen losses (NH_3 volatilization, N_2O emissions, and N leaching) from croplands: Options and mitigation strength in a global perspective. *Global Change Biology*, *25*(6), 2077–2093. https://doi.org/10.1111/gcb.14613

Liu, Y., Wu, L., Baddeley, J. A., & Watson, C. A. (2011). Models of biological nitrogen fixation of legumes. A review. *Agronomy for Sustainable Development*, *31*(1), 155–172. https://doi.org/10.1051/agro/2010008

Lobell, D., Bala, G., Mirin, A., Phillips, T., Maxwell, R., & Rotman, D. (2009). Regional differences in the influence of irrigation on climate. *Journal of Climate*, *22*(8), 2248–2255. https://doi.org/10.1175/2008JCLI2703.1

Luo, Z., Wang, E., & Sun, O. J. (2010). Can no-tillage stimulate carbon sequestration in agricultural soils? A meta-analysis of paired experiments. *Agriculture, Ecosystems and Environment*, *139*(1–2), 224–231. https://doi.org/10.1016/j.agee.2010.08.006

Lutz, F., Herzfeld, T., Heinke, J., Rolinski, S., Schaphoff, S., Von Bloh, W., et al. (2019). Simulating the effect of tillage practices with the global ecosystem model LPJmL (version 5.0-tillage). *Geoscientific Model Development*, *12*(6), 2419–2440. https://doi.org/10.5194/gmd-12-2419-2019

Lutz, F., Del Grosso, S., Ogle, S., Williams, S., Minoli, S., Rolinski, S., et al. (2020). The importance of management information and soil moisture representation for simulating tillage effects on N_2O emissions in LPJmL5.0-tillage. *Geoscientific Model Development*, *13*(9), 3905–3923. https://doi.org/10.5194/gmd-13-3905-2020

Macduff, J. H., Jarvis, S. C., & Davidson, I. A. (1996). Inhibition of N2 fixation by white clover (Trifolium repens L.) at low concentrations of NO_3^- in flowing solution culture. *Plant and Soil*, *180*(2), 287–295. https://doi.org/10.1007/BF00015312

Maillard, É., & Angers, D. A. (2014). Animal manure application and soil organic carbon stocks: A meta-analysis. *Global Change Biology*, *20*(2), 666–679. https://doi.org/10.1111/gcb.12438

Marcillo, G. S., & Miguez, F. E. (2017). Corn yield response to winter cover crops: An updated meta-analysis. *Journal of Soil and Water Conservation*, *72*(3), 226–239. https://doi.org/10.2489/jswc.72.3.226

Marino, D., Frendo, P., Ladrera, R., Zabalza, A., Puppo, A., Arrese-Igor, C., & Gonzalez, E. M. (2007). Nitrogen Fixation Control under Drought Stress. Localized or Systemic? *Plant Physiology*, *143*(4), 1968–1974. https://doi.org/10.1104/pp.106.097139

Martinsen, V., Munera-Echeverri, J. L., Obia, A., Cornelissen, G., & Mulder, J. (2019). Significant build-up of soil organic carbon under climate-smart conservation farming in Sub-Saharan Acrisols. *Science of the Total Environment*, *660*(3), 97–104. https://doi.org/10.1016/j.scitotenv.2018.12.452

Mazzoncini, M., Sapkota, T. B., Bàrberi, P., Antichi, D., & Risaliti, R. (2011). Long-term effect of tillage, nitrogen fertilization and cover crops on soil organic carbon and total nitrogen content. *Soil and Tillage Research*, *114*(2), 165–174. https://doi.org/10.1016/j.still.2011.05.001

McClelland, S. C., Paustian, K., & Schipanski, M. E. (2021). Management of cover crops in temperate climates influences soil organic carbon stocks: a meta-analysis. *Ecological Applications*, *31*(3), 1–19. https://doi.org/10.1002/eap.2278

McDermid, S. S., Mearns, L. O., & Ruane, A. C. (2017). Representing agriculture in Earth System Models: Approaches and priorities for development. *Journal of Advances in Modeling Earth Systems*, *9*(5), 2230–2265. https://doi.org/10.1002/2016MS000749

McDonagh, J. F., Toomsan, B., Limpinuntana, V., & Giller, K. E. (1995). Grain legumes and green manures as pre-rice crops in Northeast Thailand : I . Legume N_2- fixation , production and residual nitrogen benefits to rice. *Plant and Soil*, *177*, 111–126. https://doi.org/10.1007/BF00010342

Meena, B. L., Fagodiya, R. K., Prajapat, K., Dotaniya, M. L., Kaledhonkar, M. J., Sharma, P. C., et al. (2018). Legume Green Manuring: An Option for Soil Sustainability. In R. S. Meena, A. Das, G. Singh, & R. Lal (Eds.), *Legumes for Soil Health and Sustainable Management* (pp. 387–408). Springer Singapore. https://doi.org/10.1007/978-981-13-0253-4

Meinshausen, M., Nicholls, Z. R. J., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., et al. (2020). The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. *Geoscientific Model Development*, *13*(8), 3571–3605. https://doi.org/10.5194/gmd-13-3571-2020

Meyerholt, J., Zaehle, S., & Smith, M. J. (2016). Variability of projected terrestrial biosphere responses to elevated levels of atmospheric CO_2 due to uncertainty in biological nitrogen fixation. *Biogeosciences*, *13*(5), 1491–1518. https://doi.org/10.5194/bg-13-1491-2016

Michaelson-Yeates, T. P. T., Macduff, J. H., Abberton, M. T., & Raistrick, N. (1998). Characterization of novel inbred lines of white clover (Trifolium repens L.). II. Variation in N_2 fixation, NO_3^- uptake and their interactions. *Euphytica*, 45–54. https://doi.org/10.1023/A:1018369919474

Mildrexler, D. J., Zhao, M., & Running, S. W. (2011). A global comparison between station air temperatures and MODIS land surface temperatures reveals the cooling role of forests. *Journal of Geophysical Research: Biogeosciences*, *116*(3), 1–15. https://doi.org/10.1029/2010JG001486

Mínguez, M. I., Ruiz-Nogueira, B., & Sau, F. (1993). Faba bean productivity and optimum canopy development under a Mediterranean climate. *Field Crops Research*, *33*(4), 435–447. https://doi.org/10.1016/0378-4290(93)90164-I

Moebius-Clune, B. N., van Es, H. M., Idowu, O. J., Schindelbeck, R. R., Kimetu, J. M., Ngoze, S., et al. (2011). Long-term soil quality degradation along a cultivation chronosequence in western Kenya. *Agriculture, Ecosystems and Environment*, *141*(1–2), 86–99. https://doi.org/10.1016/j.agee.2011.02.018

Moss, B. (2008). Water pollution by agriculture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1491), 659–666. https://doi.org/10.1098/rstb.2007.2176

Mourtzinis, S., Kaur, G., Orlowski, J. M., Shapiro, C. A., Lee, C. D., Wortmann, C., et al. (2018). Soybean response to nitrogen application across the United States: A synthesis-analysis. *Field Crops Research*, *215*, 74–82. https://doi.org/10.1016/j.fcr.2017.09.035

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

Newbold, T. (2018). Future effects of climate and land-use change on terrestrial vertebrate community diversity under different scenarios. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1881). https://doi.org/10.1098/rspb.2018.0792

Newbold, T., Hudson, L. N., Hill, S. L. L., Contu, S., Lysenko, I., Senior, R. A., et al. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*, *520*(7545), 45–50. https://doi.org/10.1038/nature14324

Nouri, A., Lukas, S., Singh, S., Singh, S., & Machado, S. (2022). When Do Cover Crops Reduce Nitrate Leaching? A Global Meta-Analysis. *Global Change Biology*, *28*, 4736–4749. https://doi.org/10.1111/gcb.16269

Nyawira, S. S., Hartman, M. D., Nguyen, T. H., Margenot, A. J., Kihara, J., Paul, B. K., et al. (2021). Simulating soil organic carbon in maize-based systems under improved agronomic management in Western Kenya. *Soil and Tillage Research*, *211*, 105000. https://doi.org/10.1016/j.still.2021.105000

Okosun, O. O., Allen, K. C., Glover, J. P., & Reddy, G. V. P. (2021). Biology, Ecology, and Management of Key Sorghum Insect Pests. *Journal of Integrated Pest Management*, *12*(1). https://doi.org/10.1093/jipm/pmaa027

Olin, S., Schurgers, G., Lindeskog, M., Wårlind, D., Smith, B., Bodin, P., et al. (2015b). Modelling the response of yields and tissue C : N to changes in atmospheric CO_2 and N management in the main wheat regions of western Europe. *Biogeosciences*, *12*(8), 2489–2515. https://doi.org/10.5194/bg-12-2489-2015

Olin, S., Lindeskog, M., Pugh, T. A. M., Schurgers, G., Wårlind, D., Mishurov, M., et al. (2015a). Soil carbon management in large-scale Earth system modelling: Implications for crop yields and nitrogen leaching. *Earth System Dynamics*, *6*(2), 745–768. https://doi.org/10.5194/esd-6-745-2015

Osaki, M. (1993). Carbon-nitrogen interaction model in field crop production. *Plant and Soil*, *155–156*(1), 203–206. https://doi.org/10.1007/BF00025019

Osborne, S. L., & Riedell, W. E. (2011). Impact of low rates of nitrogen applied at planting on soybean nitrogen fixation. *Journal of Plant Nutrition*, *34*(3), 436–448. https://doi.org/10.1080/01904167.2011.536883

Parton, W. J., Scurlock, J. M. O., Ojima, D. S., Gilmanov, T. G., Scholes, R. J., Schimel, D. S., et al. (1993). Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide. *Global Biogeochemical Cycles*, *7*(4), 785–809. https://doi.org/10.1029/93GB02042

Patterson, T. G., & Larue, T. A. (1983). Root Respiration Associated with Nitrogenase Activity (C$_2$H$_2$) of Soybean, and a Comparison of Estimates. *Plant Physiology*, *72*(3), 701–705. https://doi.org/10.1104/pp.72.3.701

Peng, S. S., Piao, S., Zeng, Z., Ciais, P., Zhou, L., Li, L. Z. X., et al. (2014). Afforestation in China cools local land surface temperature. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(8), 2915–2919. https://doi.org/10.1073/pnas.1315126111

Penning de Vries, F. W. T., Jansen, D. M., Berge, H. F. M. ten, & Bakema, A. (1989). *Simulation of ecophysiological processes of growth in several annual crops*. Wageningen: Centre for Agricultural Publishing and Documentation (Pudoc).

Peoples, M. B., Brockwell, J., Herridge, D. F., Rochester, I. J., Alves, B. J. R., Urquiaga, S., et al. (2009). The contributions of nitrogen-fixing crop legumes to the productivity of agricultural systems. *Symbiosis*, *48*(1–3), 1–17. https://doi.org/10.1007/BF03179980

Peoples, M. B., Giller, K. E., Jensen, E. S., & Herridge, D. F. (2021). Quantifying country-to-global scale nitrogen fixation for grain legumes: I. Reliance on nitrogen fixation of soybean, groundnut and pulses. *Plant and Soil*, *469*(1–2), 1–14. https://doi.org/10.1007/s11104-021-05167-6

Pittelkow, C. M., Linquist, B. A., Lundy, M. E., Liang, X., van Groenigen, K. J., Lee, J., et al. (2015). When does no-till yield more? A global meta-analysis. *Field Crops Research*, *183*, 156–168. https://doi.org/10.1016/j.fcr.2015.07.020

Poeplau, C., & Don, A. (2015). Carbon sequestration in agricultural soils via cultivation of cover crops - A meta-analysis. *Agriculture, Ecosystems and Environment*, *200*, 33–41. https://doi.org/10.1016/j.agee.2014.10.024

Poeplau, C., Don, A., Vesterdal, L., Leifeld, J., Van Wesemael, B., Schumacher, J., & Gensior, A. (2011). Temporal dynamics of soil organic carbon after land-use change in the temperate zone - carbon response functions as a model approach. *Global Change Biology*, *17*(7), 2415–2427. https://doi.org/10.1111/j.1365-2486.2011.02408.x

Pongratz, J., Dolman, H., Don, A., Erb, K. H., Fuchs, R., Herold, M., et al. (2018). Models meet data: Challenges and opportunities in implementing land management in Earth system models. *Global Change Biology*, *24*(4), 1470–1487. https://doi.org/10.1111/gcb.13988

Portmann, F. T., Siebert, S., & Döll, P. (2010). MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, *24*(1), 1–24. https://doi.org/10.1029/2008gb003435

Porwollik, V., Rolinski, S., Heinke, J., & Müller, C. (2019). Generating a rule-based global gridded tillage dataset. *Earth System Science Data*, *11*, 823–843. https://doi.org/10.5194/essd-11-823-2019

Porwollik, V., Rolinski, S., Heinke, J., von Bloh, W., Schaphoff, S., & Müller, C. (2022). The role of cover crops for cropland soil carbon, nitrogen leaching, and agricultural yields – a global simulation study with LPJmL (V. 5.0-tillage-cc). *Biogeosciences*, *19*(3), 957–977. https://doi.org/10.5194/bg-19-957-2022

Powlson, D. S., Stirling, C. M., Jat, M. L., Gerard, B. G., Palm, C. A., Sanchez, P. A., & Cassman, K. G. (2014). Limited potential of no-till agriculture for climate change mitigation. *Nature Climate Change*, *4*(8), 678–683. https://doi.org/10.1038/nclimate2292

Powlson, D. S., Stirling, C. M., Thierfelder, C., White, R. P., & Jat, M. L. (2016). Does conservation agriculture deliver climate change mitigation through soil carbon sequestration in tropical agro-ecosystems? *Agriculture, Ecosystems and Environment*, *220*, 164–174. https://doi.org/10.1016/j.agee.2016.01.005

Prather, M. J., Hsu, J., Deluca, N. M., Jackman, C. H., Oman, L. D., Douglass, A. R., et al. (2015). Measuring and modeling the lifetime of nitrous oxide including its variability. *Journal of Geophysical Research: Atmospheres*, *120*, 5693–5705. https://doi.org/10.1002/2015JD023267

Pugh, T. A. M., Arneth, A., Olin, S., Ahlström, A., Bayer, A. D., Klein Goldewijk, K., et al. (2015). Simulated carbon emissions from land-use change are substantially enhanced by accounting for agricultural management. *Environmental Research Letters*, *10*(12). https://doi.org/10.1088/1748-9326/10/12/124008

Purcell, L. C., & Sinclair, T. R. (1990). Nitrogenase Activity and Nodule Gas Permeability Response to Rhizospheric NH_3 in Soybean. *Plant Physiology*, *92*(1), 268–272.

Quemada, M., Baranski, M., Nobel-de Lange, M. N. J., Vallejo, A., & Cooper, J. M. (2013). Meta-analysis of strategies to control nitrate leaching in irrigated agricultural systems and their effects on crop yield. *Agriculture, Ecosystems and Environment*, *174*, 1–10. https://doi.org/10.1016/j.agee.2013.04.018

Rao, M. R., & Mathuva, M. N. (2000). Legumes for improving maize yields and income in semi-arid Kenya. *Agriculture, Ecosystems and Environment*, *78*(2), 123–137. https://doi.org/10.1016/S0167-8809(99)00125-5

Rasse, D. P., Rumpel, C., & Dignac, M.-F. (2005). Is soil carbon mostly root carbon? Mechanisms for a specific stabilisation. *Plant and Soil*, *269*, 341–356. https://doi.org/10.1007/s11104-004-0907-y

Ren, W., Banger, K., Tao, B., Yang, J., Huang, Y., & Tian, H. (2020). Global pattern and change of cropland soil organic carbon during 1901-2010: Roles of climate, atmospheric chemistry,

land use and management. *Geography and Sustainability*, *1*(1), 59–69.
https://doi.org/10.1016/j.geosus.2020.03.001

Revell, L. E., Tummon, F., Salawitch, R. J., Stenke, A., & Peter, T. (2015). The changing ozone depletion potential of N$_2$O in a future climate. *Geophysical Research Letters*, *42*(22), 10047–10055. https://doi.org/10.1002/2015GL065702

Rice, W. A., Clayton, G. W., Olsen, P. E., & Lupwayi, N. Z. (2000). Rhizobial inoculant formulations and soil ph influence field pea nodulation and nitrogen fixation. *Canadian Journal of Soil Science*, *80*(3), 395–400. https://doi.org/10.4141/S99-059

Richards, F. J. (1959). A flexible growth function for empirical use. *Journal of Experimental Botany*, *10*(2), 290–301. https://doi.org/10.1093/jxb/10.2.290

Rinnofner, T., Friedel, J. K., De Kruijff, R., Pietsch, G., & Freyer, B. (2008). Effect of catch crops on N dynamics and following crops in organic farming. *Agronomy for Sustainable Development*, *28*(4), 551–558. https://doi.org/10.1051/agro:2008028

Robertson, M. J., Carberry, P. S., Huth, N. I., Turpin, J. E., Probert, M. E., Poulton, P. L., et al. (2002). Simulation of growth and development of diverse legume species in APSIM. *Australian Journal of Agricultural Research*, *53*(4), 429–446. https://doi.org/10.1071/AR01106

Rosa, L. (2022). Adapting agriculture to climate change via sustainable irrigation: Biophysical potentials and feedbacks. *Environmental Research Letters*, *17*(6). https://doi.org/10.1088/1748-9326/ac7408

Le Roux, M. R., Khan, S., & Valentine, A. J. (2009). Nitrogen and carbon costs of soybean and lupin root systems during phosphate starvation. *Symbiosis*, *48*(1–3), 102–109. https://doi.org/10.1007/BF03179989

Sacks, W. J., Cook, B. I., Buenning, N., Levis, S., & Helkowski, J. H. (2009). Effects of global irrigation on the near-surface climate. *Climate Dynamics*, *33*(2–3), 159–175. https://doi.org/10.1007/s00382-008-0445-z

Sainju, U. M., Singh, B. P., & Whitehead, W. F. (2002). Long-term effects of tillage, cover crops, and nitrogen fertilization on organic carbon and nitrogen concentrations in sandy loam soils in Georgia, USA. *Soil and Tillage Research*, *63*(3–4), 167–179. https://doi.org/10.1016/S0167-1987(01)00244-6

Sainju, U. M., Allen, B. L., Lenssen, A. W., & Ghimire, R. P. (2017). Root biomass , root/shoot ratio , and soil water content under perennial grasses with different nitrogen rates. *Field Crops Research*, *210*, 183–191. https://doi.org/10.1016/j.fcr.2017.05.029

Salvagiotti, F., Cassman, K. G., Specht, J. E., Walters, D. T., Weiss, A., & Dobermann, A. (2008). Nitrogen uptake, fixation and response to fertilizer N in soybeans: A review. *Field Crops Research*, *108*(1), 1–13. https://doi.org/10.1016/j.fcr.2008.03.001

Salzman, J., Bennett, G., Carroll, N., Goldstein, A., & Jenkins, M. (2018). The global status and trends of Payments for Ecosystem Services. *Nature Sustainability*, *1*(3), 136–144. https://doi.org/10.1038/s41893-018-0033-0

Samir, K. C., & Lutz, W. (2017). The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100. *Global Environmental Change*, *42*, 181–192. https://doi.org/10.1016/j.gloenvcha.2014.06.004

Sanderman, J., Hengl, T., & Fiske, G. J. (2017). Soil carbon debt of 12,000 years of human land use. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(7), 9575–9580. https://doi.org/10.1073/pnas.1800925115

Shen, H., Chen, Y., Hu, Y., Ran, L., Lam, S. K., Pavur, G. K., et al. (2020). Intense Warming Will Significantly Increase Cropland Ammonia Volatilization Threatening Food Security and Ecosystem Health. *One Earth*, *3*, 126–134. https://doi.org/10.1016/j.oneear.2020.06.015

Singh, M., Wanjari, R. H., Dwivedi, A., & Dalal, R. (2012). Yield response to applied nutrients and estimates of N2 fixation in 33-year-old soybean-wheat experiment on a vertisol. *Experimental Agriculture*, *48*(3), 311–325. https://doi.org/10.1017/S0014479712000129

Sleeter, B. M., Loveland, T., Domke, G., Herold, N., Wickham, J., & Wood, N. (2018). Land Cover and Land-Use Change. In D. R. Reidmiller, C. W. Avery, D. R. Easterling, K. E. Kunkel, K. L. M. Lewis, T. K. Maycock, & B. C. Stewart (Eds.), *Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment, Volume II* (pp. 202–231). Washington, DC, USA. https://doi.org/10.7930/NCA4.2018.CH5

Smil, V. (1999). Nitrogen in crop production : An account of global flows. *Global Biogeochemical Cycles*, *13*(2), 647–662. https://doi.org/10.1029/1999GB900015

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., & Zaehle, S. (2014). Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences*, *11*(7), 2027–2054. https://doi.org/10.5194/bg-11-2027-2014

Smith, P., House, J. I., Bustamante, M., Sobocká, J., Harper, R., Pan, G., et al. (2016). Global change pressures on soils from land use and management. *Global Change Biology*, *22*(3), 1008–1028. https://doi.org/10.1111/gcb.13068

Smith, P., Calvin, K., Nkem, J., Campbell, D., Cherubini, F., Grassi, G., et al. (2020). Which practices co-deliver food security, climate change mitigation and adaptation, and combat land degradation and desertification? *Global Change Biology*, *26*(3), 1532–1575. https://doi.org/10.1111/gcb.14878

Sommer, R., & Bossio, D. (2014). Dynamics and climate change mitigation potential of soil organic carbon sequestration. *Journal of Environmental Management*, *144*, 83–87. https://doi.org/10.1016/j.jenvman.2014.05.017

Sommer, R., Paul, B. K., Mukalama, J., & Kihara, J. (2018). Reducing losses but failing to sequester carbon in soils – the case of Conservation Agriculture and Integrated Soil Fertility Management in the humid tropical agro-ecosystem of Western Kenya. *Agriculture, Ecosystems and Environment*, *254*, 82–91. https://doi.org/10.1016/j.agee.2017.11.004

Song, X. P., Hansen, M. C., Potapov, P., Adusei, B., Pickering, J., Adami, M., et al. (2021). Massive soybean expansion in South America since 2000 and implications for conservation. *Nature Sustainability*, *4*(9), 784–792. https://doi.org/10.1038/s41893-021-00729-z

Soussana, J. F., Minchin, F. R., Macduff, J. H., Raistrick, N., Abberton, M. T., & Michaelson-Yeates, T. P. T. (2002). A simple model of feedback regulation for nitrate uptake and N2 fixation in contrasting phenotypes of white clover. *Annals of Botany*, *90*(1), 139–147. https://doi.org/10.1093/aob/mcf161

Stathers, T., Holcroft, D., Kitinoja, L., Mvumi, B. M., English, A., Omotilewa, O., et al. (2020). A scoping review of interventions for crop postharvest loss reduction in sub-Saharan Africa and South Asia. *Nature Sustainability*, *3*(10), 821–835. https://doi.org/10.1038/s41893-020-00622-1

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., et al. (2015). Planetary boundaries: Guiding human development on a changing planet. *Science*, *347*(6223). https://doi.org/10.1126/science.1259855

Stevenson, J. R., Serraj, R., & Cassman, K. G. (2014). Evaluating conservation agriculture for small-scale farmers in Sub-Saharan Africa and South Asia. *Agriculture, Ecosystems and Environment*, *187*, 1–10. https://doi.org/10.1016/j.agee.2014.01.018

Sutton, M. A., Oenema, O., Erisman, J., Leip, A., van Grinsven, H., & Winiwarter, W. (2011). Too much of a good thing. *Nature*, *472*, 159–161. https://doi.org/10.1038/472159a

Tewari, K., Suganuma, T., Fujikake, H., Ohtake, N., Sueyoshi, K., Takahashi, Y., & Ohyama, T. (2004). Effect of Deep Placement of N Fertilizers and Different Inoculation Methods of Bradyrhizobia on Growth, N_2 Fixation Activity and N Absorption Rate of Field-grown Soybean Plants. *Journal of Agronomy and Crop Science*, *190*(1), 46–58. https://doi.org/10.1046/j.0931-2250.2003.00073.x

Thapa, R., Mirsky, S. B., & Tully, K. L. (2018). Cover Crops Reduce Nitrate Leaching in Agroecosystems:A Global Meta-Analysis. *Journal of Environmental Quality*, *47*(6), 1400–1411. https://doi.org/10.2134/jeq2018.03.0107

Thierfelder, C., Mwila, M., & Rusinamhodzi, L. (2013). Conservation agriculture in eastern and southern provinces of Zambia: Long-term effects on soil quality and maize productivity. *Soil and Tillage Research*, *126*, 246–258. https://doi.org/10.1016/j.still.2012.09.002

Thornley, J. H. M., & Cannell, M. G. R. (2000). Modelling the components of plant respiration: Representation and realism. *Annals of Botany*, *85*(1), 55–67. https://doi.org/10.1006/anbo.1999.0997

Tian, H., Lu, C., Melillo, J., Ren, W., Huang, Y., Xu, X., et al. (2012). Food benefit and climate warming potential of nitrogen fertilizer uses in China. *Environmental Research Letters*, *7*(4). https://doi.org/10.1088/1748-9326/7/4/044020

Tian, H., Yang, J., Lu, C., Xu, R., Canadell, J. G., Jackson, R. B., et al. (2018). The global N$_2$O model intercomparison project. *Bulletin of the American Meteorological Society*, *99*(6), 1231–1251. https://doi.org/10.1175/BAMS-D-17-0212.1

Tian, H., Xu, R., Canadell, J. G., Thompson, R. L., Winiwarter, W., Suntharalingam, P., et al. (2020). A comprehensive quantification of global nitrous oxide sources and sinks. *Nature*, *586*(7828), 248–256. https://doi.org/10.1038/s41586-020-2780-0

Tian, H., Bian, Z., Shi, H., Qin, X., Pan, N., Lu, C., et al. (2022). History of anthropogenic Nitrogen inputs (HaNi) to the terrestrial biosphere: a 5 arcmin resolution annual dataset from 1860 to 2019. *Earth System Science Data*, *14*(10), 4551–4568. https://doi.org/10.5194/essd-14-4551-2022

Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(50), 20260–20264. https://doi.org/10.1073/pnas.1116437108

Tonitto, C., David, M. B., & Drinkwater, L. E. (2006). Replacing bare fallows with cover crops in fertilizer-intensive cropping systems: A meta-analysis of crop yield and N dynamics. *Agriculture, Ecosystems and Environment*, *112*(1), 58–72. https://doi.org/10.1016/j.agee.2005.07.003

Ulzen, J., Abaidoo, R. C., Mensah, N. E., Masso, C., & AbdelGadir, A. A. H. (2016). Bradyrhizobium inoculants enhance grain yields of soybean and cowpea in Northern Ghana. *Frontiers in Plant Science*, *7*, 1–9. https://doi.org/10.3389/fpls.2016.01770

Unkovich, M. J., Baldock, J., & Peoples, M. B. (2010). Prospects and problems of simple linear models for estimating symbiotic N2 fixation by crop and pasture legumes. *Plant and Soil*, *329*(1), 75–89. https://doi.org/10.1007/s11104-009-0136-5

Unkovich, Murray J, & Pate, J. S. (2000). An appraisal of recent field measurements of symbiotic N2 fixation by annual legumes. *Field Crops Research*, *65*, 211–228.

Valkama, E., Lemola, R., Känkänen, H., & Turtola, E. (2015). Meta-analysis of the effects of undersown catch crops on nitrogen leaching loss and grain yields in the Nordic countries. *Agriculture, Ecosystems and Environment*, *203*, 93–101. https://doi.org/10.1016/j.agee.2015.01.023

Vanlauwe, B., Hungria, M., Kanampiu, F., & Giller, K. E. (2019). The role of legumes in the sustainable intensification of African smallholder agriculture: Lessons learnt and challenges for the future. *Agriculture, Ecosystems and Environment*, *284*, 106583. https://doi.org/10.1016/j.agee.2019.106583

Vira, J., Hess, P., Ossohou, M., & Galy-Lacaux, C. (2022). Evaluation of interactive and prescribed agricultural ammonia emissions for simulating atmospheric composition in CAM-chem. *Atmospheric Chemistry and Physics*, *22*, 1883–1904. https://doi.org/10.5194/acp-22-1883-2022

Vitousek, P. M., Menge, D. N. L., Reed, S. C., & Cleveland, C. C. (2013). Biological nitrogen fixation: Rates, patterns and ecological controls in terrestrial ecosystems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1621). https://doi.org/10.1098/rstb.2013.0119

Voisin, A. S., Salon, C., Jeudy, C., & Warembourg, F. R. (2003). Symbiotic N2 fixation activity in relation to C economy of Pisum sativum L. as a function of plant phenology. *Journal of Experimental Botany*, *54*(393), 2733–2744. https://doi.org/10.1093/jxb/erg290

Voisin, A. S., Bourion, V., Duc, G., & Salon, C. (2007). Using an ecophysiological analysis to dissect genetic variability and to propose an ideotype for nitrogen nutrition in pea. *Annals of Botany*, *100*(7), 1525–1536. https://doi.org/10.1093/aob/mcm241

Volkholz, J., & Müller, C. (2020). ISIMIP3 soil input data (v1.0) [Data set]. https://doi.org/10.48364/ISIMIP.942125

Waha, K., Van Bussel, L. G. J., Müller, C., & Bondeau, A. (2012). Climate-driven simulation of global crop sowing dates. *Global Ecology and Biogeography*, *21*(2), 247–259. https://doi.org/10.1111/j.1466-8238.2011.00678.x

Wang, E., & Engel, T. (1998). Simulation of phenological development of wheat crops. *Agricultural Systems*, *58*(1), 1–24. https://doi.org/10.1016/S0308-521X(98)00028-6

Wang, Y., & Houlton, B. Z. (2009). Nitrogen constraints on terrestrial carbon uptake: Implications for the global carbon-climate feedback. *Geophysical Research Letters*, *36*(24), 1–5. https://doi.org/10.1029/2009GL041009

Wårlind, D., Smith, B., Hickler, T., & Arneth, A. (2014). Nitrogen feedbacks increase future terrestrial ecosystem carbon uptake in an individual-based dynamic vegetation model. *Biogeosciences*, *11*(21), 6131–6146. https://doi.org/10.5194/bg-11-6131-2014

West, P. C., Gerber, J. S., Engstrom, P. M., Mueller, N. D., Brauman, K. A., Carlson, K. M., et al. (2014). Leverage points for improving global food security and the environment. *Science*, *345*(6194), 325–328. https://doi.org/10.1126/science.1246067

Wickham, J., Wade, T. G., & Riitters, K. H. (2014). An isoline separating relatively warm from relatively cool wintertime forest surface temperatures for the southeastern United States. *Global and Planetary Change*, *120*, 46–53. https://doi.org/10.1016/j.gloplacha.2014.05.012

Williams, C. M., King, J. R., Ross, S. M., Olson, M. A., Hoy, C. F., & Lopetinsky, K. J. (2014). Effects of three pulse crops on subsequent barley, canola, and wheat. *Agronomy Journal*, *106*(2), 343–350. https://doi.org/10.2134/agronj2013.0274

Wittwer, R. A., Dorn, B., Jossi, W., & Van Der Heijden, M. G. A. (2017). Cover crops support ecological intensification of arable cropping systems. *Scientific Reports*, *7*, 41911. https://doi.org/10.1038/srep41911

Wortmann, C., Mamo, M., Mburu, C., Letayo, E., Birru, G. A., Kaizzi, K. C., et al. (2009). *Atlas of Sorghum Production in Eastern and Southern Africa*. Lincoln, United States.

Wu, L., & McGechan, M. B. (1999). Simulation of nitrogen uptake, fixation and leaching in a grass/white clover mixture. *Grass and Forage Science*, *54*(1), 30–41. https://doi.org/10.1046/j.1365-2494.1999.00145.x

Wu, L., Misselbrook, T. H., Feng, L., & Wu, L. (2020). Assessment of nitrogen uptake and biological nitrogen fixation responses of soybean to nitrogen fertiliser with SPACSYS. *Sustainability*, *12*(15). https://doi.org/10.3390/SU12155921

Xia, L., & Yan, X. (2023). How to feed the world with less nitrogen pollution. *Nature*, *613*, 34–35. https://doi.org/10.1038/d41586-022-04490-x

Xia, L., Lam, S. K., Wolf, B., Kiese, R., Chen, D., & Butterbach-Bahl, K. (2018). Trade-offs between soil carbon sequestration and reactive nitrogen losses under straw return in global agroecosystems. *Global Change Biology*, *24*(12), 5919–5932. https://doi.org/10.1111/gcb.14466

Xu-Ri, & Prentice, I. C. (2017). Modelling the demand for new nitrogen fixation by terrestrial ecosystems. *Biogeosciences*, *14*(7), 2003–2017. https://doi.org/10.5194/bg-14-2003-2017

Yu, M., Gao, Q., & Shaffer, M. J. (2002). Simulating Interactive Effects of Symbiotic Nitrogen Fixation, Carbon Dioxide Elevation, and Climatic Change on Legume Growth. *Journal of Environmental Quality*, *31*(2), 634–641. https://doi.org/10.2134/jeq2002.6340

Zaehle, S., & Friend, A. D. (2010). Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates. *Global Biogeochemical Cycles*, *24*(1), 1–13. https://doi.org/10.1029/2009GB003521

Zaehle, S., Ciais, P., Friend, A. D., & Prieur, V. (2011). Carbon benefits of anthropogenic reactive nitrogen offset by nitrous oxide emissions. *Nature Geoscience*, *4*(9), 601–605. https://doi.org/10.1038/ngeo1207

Zahran, H. H. (1999). Rhizobium-Legume Symbiosis and Nitrogen Fixation under Severe Conditions and in an Arid Climate. *Microbiology and Molecular Biology Reviews*, *63*(4), 968–989. https://doi.org/10.1128/mmbr.63.4.968-989.1999

Zhang, B., Tian, H., Lu, C., Dangal, S., Yang, J., & Pan, S. (2017). Manure nitrogen production and application in cropland and rangeland during 1860–2014: A 5-minute gridded global data set for Earth system modeling. *Earth System Science Data*, *9*, 667–678. https://doi.org/10.5194/essd-2017-11

Zhang, J., Tian, H., Yang, J., & Pan, S. (2018). Improving Representation of Crop Growth and Yield in the Dynamic Land Ecosystem Model and Its Application to China. *Journal of*

Advances in Modeling Earth Systems, *10*(7), 1680–1707.
https://doi.org/10.1029/2017MS001253

Zhang, X., Zou, T., Lassaletta, L., Mueller, N. D., Tubiello, F. N., Lisk, M. D., et al. (2021).
Quantification of global and national nitrogen budgets for crop production. *Nature Food*, *2*(7),
529–540. https://doi.org/10.1038/s43016-021-00318-5

Zhang, X., Davidson, E. A., Mauzerall, D. L., Searchinger, T. D., Dumas, P., & Shen, Y. (2015).
Managing nitrogen for sustainable development. *Nature*, *528*(7580), 51–59.
https://doi.org/10.1038/nature15743

Zhao, J., Chen, J., Beillouin, D., Lambers, H., Yang, Y., Smith, P., et al. (2022). Global systematic
review with meta-analysis reveals yield advantage of legume-based rotations and its drivers.
Nature Communications, *13*, 4926. https://doi.org/10.1038/s41467-022-32464-0

Zhao, K., & Jackson, R. B. (2014). Biophysical forcings of land-use changes from potential forestry
activities in North America. *Ecological Monographs*, *84*(2), 329–353.
https://doi.org/10.1890/12-1705.1

Zomer, R. J., Bossio, D. A., Sommer, R., & Verchot, L. V. (2017). Global Sequestration Potential
of Increased Organic Carbon in Cropland Soils. *Scientific Reports*, *7*, 15554.
https://doi.org/10.1038/s41598-017-15794-8

Appendix

The original published papers are attached in the following.

Geoscientific
Model Development

Modeling symbiotic biological nitrogen fixation in grain legumes globally with LPJ-GUESS (v4.0, r10285)

Jianyong Ma[1], **Stefan Olin**[2], **Peter Anthoni**[1], **Sam S. Rabin**[1], **Anita D. Bayer**[1], **Sylvia S. Nyawira**[3], and **Almut Arneth**[1,4]

[1]Institute of Meteorology and Climate Research-Atmospheric Environmental Research, Karlsruhe Institute of Technology, 82467 Garmisch-Partenkirchen, Germany
[2]Department of Physical Geography and Ecosystems Science, Lund University, 22362 Lund, Sweden
[3]International Center for Tropical Agriculture (CIAT), ICIPE Duduville Campus, P.O. Box 823-00621, Nairobi, Kenya
[4]Institute of Geography and Geoecology, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany

Correspondence: Jianyong Ma (jianyong.ma@kit.edu)

Abstract. Biological nitrogen fixation (BNF) from grain legumes is of significant importance in global agricultural ecosystems. Crops with BNF capability are expected to support the need to increase food production while reducing nitrogen (N) fertilizer input for agricultural sustainability, but quantification of N fixing rates and BNF crop yields remains inadequate on a global scale. Here we incorporate two legume crops (soybean and faba bean) with BNF into a dynamic vegetation model LPJ-GUESS (Lund–Potsdam–Jena General Ecosystem Simulator). The performance of this new implementation is evaluated against observations from a range of water and N management trials. LPJ-GUESS generally captures the observed response to these management practices for legume biomass production, soil N uptake, and N fixation, despite some deviations from observations in some cases. Globally, simulated BNF is dominated by soil moisture and temperature, as well as N fertilizer addition. Annual inputs through BNF are modeled to be $11.6 \pm 2.2\,\mathrm{Tg\,N}$ for soybean and $5.6 \pm 1.0\,\mathrm{Tg\,N}$ for all pulses, with a total fixation of $17.2 \pm 2.9\,\mathrm{Tg\,N\,yr^{-1}}$ for all grain legumes during the period 1981–2016 on a global scale. Our estimates show good agreement with some previous statistical estimates but are relatively high compared to some estimates for pulses. This study highlights the importance of accounting for legume N fixation process when modeling C–N interactions in agricultural ecosystems, particularly when it comes to accounting for the combined effects of climate and land-use change on the global terrestrial N cycle.

1 Introduction

The agricultural sector is the main contributor to anthropogenic nitrous oxide (N_2O) emissions (Reay et al., 2012; Tian et al., 2020) and a key nitrate pollution source to freshwater systems (Moss, 2008), mostly due to the intensive use of synthetic nitrogen (N) fertilizer and animal manure (Lu and Tian, 2017). This trend has been amplified by the expansion of agricultural land to provide food for a growing population and changing dietary patterns (FAO, 2018). The use of crops with biological N fixation (BNF) capability in agriculture has been discussed as one option to address the conflict between the need to increase food production and the associated environmental problems of N loss (Becker, et al., 1995; Fageria, 2007; Northup and Rao, 2016). N-fixing crops, like grain and forage legumes, not only provide protein-rich food for the human population and farmed animals (Voisin et al., 2014; Stagnari et al., 2017), but they are also directly useable as "green manure", reducing the amount of chemical N fertilizer required in agricultural systems (Liu et al., 2011; Meena et al., 2018).

Soybean (*Glycine max* L.), with its countless and varied uses, is now one of the most widely grown crops in the world because of attractive cash return from its grain yield (FAO-STAT, 2021). There are concerns about the sustainability of soybean production, in particular because of its links to deforestation and loss of native vegetation in the Amazon and other areas of South America (Fehlenberg et al., 2017; Heilmayr et al., 2020). Other grain legumes, such as faba bean

(*Vicia faba* L.), chickpea (*Cicer arietinum* L.), and cowpea (*Vigna unguiculata* L.), play an important role in improving soil quality as green manure when they are rotated or used as intercrops between cereals depending on the region (Williams et al., 2014; Denton et al., 2017). In comparison to non-legume plants, using legumes as green manure is more effective to build up or maintain soil fertility, as they not only increase soil organic matter when adding their biomass to soils, but also add extra N into the soil as a result of their symbiotic association with bacteria (Peoples et al., 2009; Ciampitti and Salvagiotti, 2018). The enriched soil N and soil organic carbon contents jointly support growth and productivity in subsequent crops (Jensen et al., 2012; Hajduk et al., 2015). Much experimental evidence has indicated that grain legume biomass increases linearly with an increasing BNF rate (Salvagiotti et al., 2008; Unkovich et al., 2010; Córdova et al., 2019) and that the N benefit to soil fertility from green manure is closely correlated with N fixation capacity, assuming that the entire legume plant is tilled into the soil (Fageria, 2007; Meena et al., 2018). Estimating the rate of BNF is thus important not only for an accurate prediction of grain legume production but also for a better understanding of where and to what degree N loss (i.e., N leaching and gaseous N emission) in cropland systems can be reduced by partially or fully replacing chemical N fertilizer with legume green manure.

Although grain legumes' BNF rates can be measured at field sites and in controlled environments, ecological models are needed for understanding and quantifying the rate of BNF on larger spatial scales and longer temporal perspectives. In many process-based crop models, a common method of representing BNF is to use a pre-defined potential or maximum N fixation rate that is adjusted by limiting environmental factors (Liu et al., 2011). The potential N fixation rate is then estimated either from plant nodule, root, and aboveground biomass (e.g., Boote et al., 2008; Corre-Hellou et al., 2009; Wu et al., 2020) or from plant N demand status (e.g., Cabelguenne et al., 1999; Robertson et al., 2002), varying with plant life cycle. Environmental constraining factors, such as soil temperature, water availability, soil mineral N concentration, and plant growth stage, are mostly taken into account (Liu et al., 2011; Chen et al., 2016). The big challenge in modeling legume BNF is that the process of symbiotic N fixation is always accompanied by the cost of fixed total photosynthetic carbon (C) to maintain legume symbiotic growth, activity, and reserves, which may be around 4 %–16 % of C (Kaschuk et al., 2009). Such a photosynthetic consumption strength would result in productivity loss if the photosynthesis rate did not increase to compensate for the cost (Kaschuk et al., 2010). In most models C cost mechanisms have not been implemented into N fixation, consistent with the assumption that the plant N uptake from soils does not cost carbon (e.g., Cabelguenne et al., 1999; Robertson et al., 2002; Corre-Hellou et al., 2009; Drewniak et al., 2013; Von Bloh et al., 2018; Wu et al., 2020), despite many field experi-

ments demonstrating that the energy consumption required for BNF is far larger than soil mineral N uptake (Ryle et al., 1979; Harris et al., 1985; Macduff et al., 1996). In several other models, root substrate C concentration was adopted as an alternative to represent the C demand of N fixation (e.g., Thornley and Cannell, 2000; Yu and Zhuang, 2020). Only a few models assume that such a consumption can be assessed directly against C acquired in photosynthesis, in which the C cost per unit of fixed N is defined as either a constant of $6\,\mathrm{kg\,C\,kg\,N^{-1}}$ (Boote et al., 2008; Meyerholt et al., 2016) or a dynamic function of soil temperature ranging between 7.5 and $12.5\,\mathrm{kg\,C\,kg\,N^{-1}}$ (Houlton et al., 2008; Fisher et al., 2010).

The global production and consumption of grain legumes have greatly increased over recent decades (FAOSTAT, 2021). Accurately representing and quantifying the dynamic process of biological N fixation in models is important for better understanding grain legumes' contribution to food security and agriculture sustainability, particularly in the context of global environmental change. However, because of inadequate information on the environment and crop management, as well as the missing or incomplete BNF mechanism in models (e.g., C cost as mentioned above), current simulation of grain legume N fixation and its yield is still very weak, especially when it comes to global-scale modeling.

Thus, in this study, by accounting for the importance of soybean in overall agriculture and trade, as well as the higher N fixation capacity of faba bean compared to other pulses (Peoples et al., 2009; Unkovich et al., 2010; Denton et al., 2017; Liu et al., 2019), we implement these two grain legumes with BNF into a process-based vegetation model (LPJ-GUESS; Smith et al., 2014; Olin et al., 2015). Processes are added to LPJ-GUESS to estimate the symbiotic relationship between legumes and bacteria, also taking into account the plant C cost of BNF. Model results are extensively evaluated with worldwide site-level observed data and compared against country-level yield statistics, as well as continent-level BNF rates. The model-based and large-scale quantification of the N fixation capacity in legumes provides a scientific foundation for predicting the present and future N cycle in agro-ecosystems, allowing recommendations for fertilizer N application under different climatic conditions in legume-based farming systems.

2 Methods

2.1 Model description

LPJ-GUESS is a process-based dynamic vegetation model that simulates carbon and nitrogen (C–N) dynamics at scales typically ranging from regional to global (Smith et al., 2014). The model represents vegetation and soil dynamic processes as well as their interactions in response to changes in the environment and management, such as climate, CO_2 concentra-

tion, soil physical properties, N deposition, and N fertilization. Three land-use types are included in the model: natural vegetation, pasture, and cropland. Vegetation on natural land is represented as the establishment, growth, and mortality of 12 plant function types (PFTs). Pastures are simulated by competing C_3 and C_4 grasses, in which 50 % of aboveground biomass is annually harvested to account for the effects of grazing (Lindeskog et al., 2013). Crops in LPJ-GUESS are described by crop functional types (CFTs), which differ in their C allocation scheme, morphological traits, and heat sum requirement for growth. At present, four CFTs are represented in the C–N version of LPJ-GUESS: two temperate C_3 crops with sowing carried out in spring and autumn, a tropical C_3 crop (representing rice), and a C_4 crop (representing maize). Sowing dates on a large scale are determined dynamically in the model based on local climatology in each grid cell with five seasonality types represented (a combination of temperature- and precipitation-limited behaviors; Waha et al., 2012), and crops are harvested once each year when prescribed heat sum requirements are fulfilled (Lindeskog et al., 2013). Multi-cropping systems within a year are not yet implemented in the model. The recent representation of crops includes the incorporation of soil N transformation (Tian et al., 2020) together with a C–N allocation for crops operating on a daily time step (Lindeskog et al., 2013; Olin et al., 2015). Cropland management options for global-scale application include irrigation, tillage, N application, cover crop grass between the main growing seasons, and residue retention (Pugh et al., 2015; Olin et al., 2015). In this study, soybean is simulated as one additional crop because of its large importance as a food, fodder, and oil crop, and the parametrization of faba bean is representative for the group of pulses in general. The model schematic and other calculations including the C cycle and the N cycle follow an earlier version of LPJ-GUESS (Smith et al., 2014; Wårlind et al., 2014; Olin et al., 2015).

2.2 Updated daily carbon allocation parameters

Similar to most ecosystem and crop models, LPJ-GUESS adopts crop-specific accumulated heat requirements to model plant growth development, and crops are allowed to adapt to the local climate by dynamically adjusting the heat requirements to different climatic zones (Lindeskog et al., 2013). To better represent C and N allocation in various phenological phases, Olin et al. (2015) defined crop development stage by considering the effects of temperature, vernalization days, and photo-period following Wang and Engel (1998). In this study, we assume that the grain legume development stage is linearly correlated with its accumulated heat units according to the field-based soybean experiments described in Irmak et al. (2013). It is estimated as

$$\text{DS} = \begin{cases} a_{\text{veg}} + b_{\text{veg}} \times fphu & (fphu \leq \text{fphu}_{\text{anthesis}}) \\ a_{\text{rep}} + b_{\text{rep}} \times fphu & (fphu > \text{fphu}_{\text{anthesis}}) \end{cases}, \quad (1)$$

where DS is crop development stage ranging from 0 to 2 (DS = 0, sowing; DS = 1, flowering; DS = 2, harvest); $fphu$ is the fraction of today's accumulated heat units to the total heat requirement; $\text{fphu}_{\text{anthesis}}$ is the threshold of $fphu$ when anthesis starts, below (above) which crop growth belongs to the vegetative (reproductive) stage; and a and b are the linear regression coefficients, varying between the vegetative and reproductive phases. The values of a and b, as well as the crop-specific base temperature (° C) to estimate the accumulated heat units, are both given in Table S1 in the Supplement.

The daily fraction of assimilate allocation to leaves, stems, and roots is an important process before storage organs are formed. The assimilate invested in roots can help crops overcome water or nutrient limitation when they suffer from stress in the vegetative stage, whereas new assimilate invested in leaves generally gives a highly efficient return from the photosynthesis product (Penning de Vries et al., 1989). Unlike cereal crops, nodulated plants, particularly soybeans, are more likely to achieve a higher photosynthesis rate and delay leaf senescence due to the continued N supply from biological N fixation (Abu-shakra et al., 1978; Kaschuk et al., 2010). A precise representation of assimilate partitioning to the plant organs when modeling BNF in grain legumes is especially important considering the high C cost from fixing N from the atmosphere. Productivity loss would be simulated if the leaf photosynthesis rate did not increase to compensate for the costs (Macduff et al., 1996; Kaschuk et al., 2009).

Following Olin et al. (2015), relationships between assimilate allocation to legume organs were established based on the data from Penning de Vries et al. (1989) and Boote et al. (2002). We fitted the allocation functions using a Richards logistic growth curve (Eq. 2, Richards, 1959) to model the allocation to each organ dynamically and separately. For each allocation function f_i (see Eqs. 3–5 below),

$$f_i = a_i + \frac{b_i - a_i}{1 + e^{-C_i \times (\text{DS} - d_i)}}, \quad (2)$$

where DS is crop development stage, and a_i, b_i, c_i, and d_i are fitting coefficients for the three functions (specific values given in Table S1).

Maintaining BNF in the reproductive stage (i.e., after anthesis; DS > 1) would reduce the flow of carbon assimilation to storage organs. We adjusted the allocation functions from Olin et al. (2015) so that the model allowed a dynamic adaptation of the allocation to grain over the seed-filling period in response to BNF cost (see Eqs. 3–5 for details).

2.2.1 Yield vs. the whole plant

After anthesis (DS > 1), most assimilates are allocated and re-translocated from the vegetative organs to the grains. During the late seed-filling period (DS $\geq d_1$, see Eq. 3), we assumed that the fraction of carbon allocated to yield would increase to partly compensate for the productivity loss caused by spending on N fixation at the cost of reducing the flow

of carbon to leaves and stem (see Eq. 4). We established the ratio of the allocation to yield relative to the whole plant as

$$f_1 = \frac{P_{\text{yield}}}{P_{\text{veg}} + P_{\text{yield}}}$$

$$= \begin{cases} a_1 + \frac{b_1 - a_1}{1 + e^{-c_1 \times (\text{DS} - d_1)}} & \text{DS} < d_1 \\ \left(a_1 + \frac{b_1 - a_1}{1 + e^{-c_1 \times (\text{DS} - d_1)}}\right) \times (1 + P_{\text{BNFcost}}) & \text{DS} \geq d_1 \end{cases}, \quad (3)$$

where P_{yield} and P_{veg} are the fraction of carbon allocated to yield and vegetative organs, respectively, ranging from 0 to 1; P_{BNFcost} is the proportion of net primary production (NPP) used for BNF to today's total NPP; and d_1 is the fitting coefficient representing the DS of the maximum growth rate of grain ($d = 1.41$ for soybean and 1.46 for faba bean, see Table S1).

2.2.2 Leaf vs. shoot vegetative organs

Similarly, the ratio of leaf vs. shoot vegetative allocation is specified as

$$f_2 = \frac{P_{\text{leaf}}}{P_{\text{veg}} - P_{\text{root}}}$$

$$= \begin{cases} a_2 + \frac{b_2 - a_2}{1 + e^{-c_2 \times (\text{DS} - d_2)}} & \text{DS} < d_1 \\ \left(a_2 + \frac{b_2 - a_2}{1 + e^{-c_2 \times (\text{DS} - d_2)}}\right) - P_{\text{BNFcost}} & \text{DS} \geq d_1 \end{cases}, \quad (4)$$

where P_{leaf} and P_{root} are the fraction of carbon allocated to leaf and root, respectively. The fitting function of leaf vs. shoot vegetative organs in soybean is given in Fig. 1a.

2.2.3 Root vs. vegetative organs

When a plant experiences water or nutrient stress, it invests more assimilate to roots relative to shoot vegetative organs (Penning de Vries et al., 1989). We implemented dynamic increases in the allocation to roots during the late seed-filling period to help legumes cope with the C loss from BNF cost and established the relationship between the allocation to root and that to vegetative organs as

$$f_3 = \frac{P_{\text{root}}}{P_{\text{veg}}}$$

$$= \begin{cases} a_3 + \frac{b_3 - a_3}{1 + e^{-c_3 \times (\text{DS} - d_3)}} & \text{DS} < d_1 \\ \left(a_3 + \frac{b_3 - a_3}{1 + e^{-c_3 \times (\text{DS} - d_3)}}\right) + (1 - f_1) \times P_{\text{BNFcost}} & \text{DS} \geq d_1 \end{cases}. \quad (5)$$

In addition, carbon partitioning to vegetative organs (P_{veg}) can be calculated by subtracting the reproductive allocation (i.e., P_{yield}) from the whole plant as

$$P_{\text{veg}} + P_{\text{yield}} = 1 \Rightarrow P_{\text{veg}} = 1 - P_{\text{yield}} = 1 - f_1. \quad (6)$$

Finally, we can achieve dynamic carbon allocation to the plant organs over the growing season by combining Eqs. (3)–(6).

$$\begin{cases} P_{\text{yield}} = f_1 \\ P_{\text{leaf}} = f_2 \times (1 - f_1) \times (1 - f_3) \\ P_{\text{stem}} = (1 - f_1) \times (1 - f_2) \times (1 - f_3) \\ P_{\text{root}} = f_3 \times (1 - f_1) \end{cases} \quad (7)$$

Partitioning functions are plotted for soybean in Fig. 1b and for faba bean in Fig. S1 in the Supplement. Significant differences in allocation patterns can exist between cultivars. Compared to cereals (Olin et al., 2015), we found that grain legumes are more likely to allocate more assimilate to leaves not only in partitioning proportion but also in the length of allocation time, probably corresponding to their higher leaf activities in response to N fixation (Kaschuk et al., 2010).

2.3 Representation of BNF

Fixing N from the atmosphere and N uptake from soils represents two N sources for grain legumes to meet their total plant N demand. The latter has a higher priority for plants because the process is less energy-consuming than N fixation (Ryle et al., 1979; Macduff et al., 1996). Following on this idea, in LPJ-GUESS, N fixation will only be triggered when the following two assumptions are valid at the same time (Fig. 2): (1) if today's plant growth still suffers from N limitation after N uptake from soils (i.e., the N deficit, plant N demand minus soil N uptake, is greater than zero). The plant will then be allowed to fix N from the atmosphere to fill the N deficit. (2) Since N fixation is strongly related to photosynthetic assimilate due to its high energy consumption, BNF in the model is assumed to take place only when today's NPP is positive so that adequate C supply can be provided to meet the BNF cost.

Modeling the BNF rate is adapted from previously published methods (e.g., CROPGRO, EPIC, APSIM; see Liu et al., 2011) in that it considers (1) the potential N fixation rate, (2) the limitation of temperature, (3) soil water status, and (4) the crop growth stage as

$$N_{\text{fix}} = N_{\text{fixpot}} \times f_{\text{T}} \times f_{\text{W}} \times f_{\text{DS}}, \quad (8)$$

where N_{fix} is the N fixation rate; N_{fixpot} is the potential N fixation rate; and f_{T}, f_{W}, and f_{DS} are limitations (ranging 0 to 1) on BNF by soil temperature, soil water availability, and crop development stage function, respectively.

The definition of the potential N fixation rate in some studies is based on the strong relationship between N fixation and either nodule size, biomass (Weisz et al., 1985; Voisin et al., 2003), or root dry matter (Soussana et al., 2002; Voisin et al., 2007). Due to the difficulties in measuring both nodules and roots in the field directly, some studies also adopt shoot biomass to replace nodule or root biomass based on the empirical relationship between these two variables (Yu et al., 2002; Corre-Hellou et al., 2009; Wu et al., 2020). In our implementation, since the nodulation process of legumes has
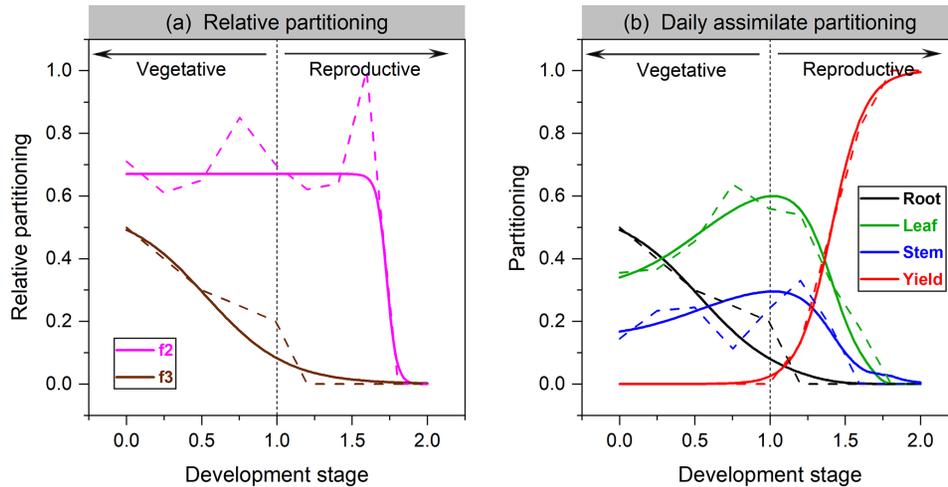
Figure 1. The organ's relative allocation **(a)** and assimilate partitioning **(b)** to roots, leaves, stem, and yields for soybean. Solid lines represent the fitted Richards functions in this study, and dashed lines are the allocation scheme from Penning de Vries et al. (1989). f_2 in **(a)** denotes leaf relative allocation to shoot vegetative organs (Eq. 4), whereas f_3 is root relative allocation to vegetative organs (Eq. 5).

not yet been implemented in LPJ-GUESS, N_{fixpot} is assumed to be proportional to root dry matter:

$$N_{\text{fixpot}} = N_{\text{maxfixpot}} \times \text{DM}_{\text{root}}, \tag{9}$$

where $N_{\text{maxfixpot}}$ is the maximum nitrogen fixation rate of roots (g N g^{-1} root DM), and DM_{root} is root dry matter (g root DM m^{-2}). Since the experimental parameter $N_{\text{maxfixpot}}$ is strongly related to the effectiveness of rhizobial strains and varies widely between species and sites, it is not easy to obtain the parameter for each legume crop. In this study, we assume that legumes are either inoculated or there are high enough populations of strains in the soil that $N_{\text{maxfixpot}}$ is not constrained by the effectiveness of rhizobia. Here $N_{\text{maxfixpot}}$ is assumed to be a constant as 0.03 g N g^{-1} root DM for both grain legumes as a moderate value taken from the literature (Soussana et al., 2002; Eckersten et al., 2006; Boote et al., 2008).

Soil temperature is a controlling factor for both microbial activities and plant growth. For soybean, 20–35 °C has been found to be optimal for nitrogenase activity and for faba bean the optimal soil temperature can range from 16–25 °C (Boote et al., 2008). The influence of soil temperature on legume BNF is represented in the model as a four-threshold-temperature function:

$$f_{\text{T}} = \begin{cases} 0 & (T < T_{\min} \text{ or } T > T_{\max}) \\ \dfrac{T - T_{\min}}{T_{\text{optL}} - T_{\min}} & (T_{\min} \leq T < T_{\text{optL}}) \\ 1 & (T_{\text{optL}} \leq T \leq T_{\text{optH}}) \\ \dfrac{T_{\max} - T}{T_{\max} - T_{\text{optH}}} & (T_{\text{optH}} < T \leq T_{\max}) \end{cases}, \tag{10}$$

where T is soil temperature (°C) at a depth of 25 cm representing the mean temperature of the topsoil layer in the

model (0–50 cm), $T_{\min}$ ($T_{\max}$) is the minimum (maximum) temperature below (above) which N fixation stops, and T_{optL} and T_{optH} are the lower and higher optimal temperatures within which N fixation is not limited by temperature. The values of these four temperature thresholds vary among legume crops and are given in Table 1.

In addition to temperature, soil water content is a major factor controlling the rate of N fixation (Srivastava and Ambasht, 1994). Too little water strongly inhibits BNF due to impacts of drought stress on nodule nitrogenase activity (Serraj et al., 1999; Marino et al., 2007). Although oxygen is needed to support the respiration of legume roots and bacteria in the nodules, nitrogenase is more active in anoxic, waterlogged environments (Jiang et al., 2021). A linear water limitation function is thus incorporated into LPJ-GUESS (Wu and McGechan, 1999) and is represented as

$$f_{\text{W}} = \begin{cases} 0 & (W_f \leq W_a) \\ \varphi_1 + \varphi_2 \times W_f & (W_a < W_f < W_b) \\ 1 & (W_f \geq W_b) \end{cases}, \tag{11}$$

where W_f is relative soil water content in the topsoil layer (0–50 cm) ranging from 0 to 1, φ_1 and φ_2 are empirical coefficients, W_a is the threshold of W_f below which N fixation is fully restricted by soil water deficit, and W_b is the value above which N fixation is not inhibited by soil water content. The values of the parameters are shown in Table 1.

The influence of plant growth stage on legume BNF rate is taken into account in very few models; the process is generally stopped forcibly after the crop reaches a certain development stage. For example, in the CROPGRO model (Boote et al., 2008), N fixation in soybean starts in the second trifoliolate stage and continues until the end of physiological maturity, whereas it ceases at the middle of the seed-filling period

in the EPIC model (Cabelguenne et al., 1999). Much experimental evidence has indicated that the N fixed by legumes varies widely among crop growth stages, with the largest BNF rate observed between the late vegetative phase and the early seed-filling period (Santachiara et al., 2017; Córdova et al., 2020; Ciampitti et al., 2021). In this study, a specific function, similar to the temperature response function, is thus implemented in the BNF scheme to represent the variation of N fixation with the course of the legume life cycle:

$$
\begin{aligned}
&f_{DS} \\
&=
\begin{cases}
0 & (\text{NDS} < \text{NDS}_{min} \text{ or } \text{NDS} > \text{NDS}_{max}) \\
\dfrac{\text{NDS} - \text{NDS}_{min}}{\text{NDS}_{optL} - \text{NDS}_{min}} & (\text{NDS}_{min} \le \text{NDS} < \text{NDS}_{optL}) \\
1 & (\text{NDS}_{optL} \le \text{NDS} \le \text{NDS}_{optH}) \\
\dfrac{\text{NDS}_{max} - \text{NDS}}{\text{NDS}_{max} - \text{NDS}_{optH}} & (\text{NDS}_{optH} < \text{NDS} \le \text{NDS}_{max})
\end{cases}
\end{aligned}
\tag{12}
$$

where NDS is normalized crop development stage ranging from 0 to 1 (0, sowing; 0.5, flowering; 1, harvest), NDS_{min} is the time before which there is no N fixation due to inadequate nodulation, NDS_{max} is the time after which N fixation suspends due to nodule senescence, and NDS_{optL} and NDS_{optH} define the period within which the legume BNF rate is not inhibited by development stage. The values of the parameters for two grain legumes are derived from the literature and listed in Table 1.

In addition to the environmental limitation factors, the amount of daily NPP also affects N fixation in the model. The NPP requirement for BNF cost is computed based on the estimated N fixation rate (N_{fix}, Eq. 8) by multiplying the C cost per unit fixed N, which is assumed to be a fixed value of $6\,\text{g C g}^{-1}$ N as a moderate value taken from previous studies (Ryle et al., 1979; Patterson and Larue, 1983; Boote et al., 2008; Kaschuk et al., 2009). The NPP cost to maintain BNF is released as CO_2 to the atmosphere and modeled as part of the autotrophic respiration of the soil (Fig. 2). Since the fixed N is partly transported to plant leaves and continues to support photosynthesis, the plant may get additional C profits from the investment of BNF by enhancing the leaf N content that optimizes the carboxylation capacity (V_{max}) (Kull, 2002). Following on this idea, another assumption adopted in this study is that at most 50 % of today's NPP can be used for N fixation before the crops reach the development stage of grain maximum growth rate (DS $< d_1$, see Eq. 13). After this the maximum proportion of today's NPP used for BNF cost is dynamically reduced and assumed to be the fraction of carbon allocation to leaves and stem:

$$
\begin{aligned}
&\text{MAXNPP}_{BNFcost} \\
&=
\begin{cases}
0.5 & \text{DS} < d_1 \\
P_{leaf} + P_{stem} = (1 - f_1) \times (1 - f_3) & \text{DS} \ge d_1
\end{cases}
\end{aligned}
\tag{13}
$$

where $\text{MAXNPP}_{BNFcost}$ is the maximum proportion of today's NPP used for N fixation varying from 0–0.5, and P_{leaf} and P_{stem} are the fraction of carbon (i.e., NPP) allocated to

leaf and stem, respectively (see Eq. 7 for details). A flowchart of the BNF scheme in LPJ-GUESS is shown in Fig. 2.

2.4   Experimental setup

Field-based data from the literature, together with global yield statistics from legume-producing countries and region-level N fixation data from published sources, were compared to model runs to examine performance in simulating yields and BNF rate from the site scale to a larger region.

In order to build up cropland soil C and N pools, all simulations were initialized with a 500-year spin-up using atmospheric CO_2 from 1901 combined with repeating detrended 1901–1930 climate from GSWP3-W5E5 (Dirmeyer et al., 2006; Lange, 2019; Cucchi et al., 2020). The cropland fraction linearly increased from zero to the first historic value (1901) during the last 30 years of spin-up. Monthly atmospheric N deposition (NH_x, NO_y) was used as simulated by CCMI (NCAR Chemistry–Climate Model Initiative). The value was interpolated to $0.5° \times 0.5°$ from the original resolution ($1.9° \times 2.5°$) to match the resolution of the climate data (Tian et al., 2018). Below, the setup of the different experiments is explained in detail.

2.4.1   Model evaluation at site scale

To evaluate the model's ability to simulate BNF rate and yields, field-based N fixation trials with detailed measurements of soil N uptake, biomass, and N mass allocation were collected from the published literature. This dataset comprised 17 soybean and 7 faba bean sites located between $\sim 33°$ S and $\sim 53°$ N (Fig. 3). In these trials, BNF response to various management practices (such as N fertilizer addition and irrigation) were investigated. Details about these sites – their geographic coordinates, BNF trials, and the years of available data, as well as corresponding site-specific plant traits (e.g., specific leaf area and grain C : N ratio) – are provided in Table S2.

In some field experiments, BNF rate and/or soil N uptake are not directly reported in the literature, so we estimated these values as

$$
\begin{cases}
\text{BNF}_{obs} = \%\text{Ndfa} \times N_{plant} \\
\text{SoilNuptake}_{obs} = (100 - \%\text{Ndfa}) \times N_{plant},
\end{cases}
\tag{14}
$$

where %Ndfa is the proportion of plant N derived from the atmosphere (ranging 0–100), representing the contribution of N fixation to the plant total N uptake, and N_{plant} is the amount of N accumulated in the plant (kg N ha^{-1}), defined as either the shoot or the whole plant N mass, depending on the measurement method adopted in the experiment.

In general, grain yields, plant tissue dry mass, and N mass, together with %Ndfa, soil N uptake, and N fixation, are widely measured variables in field-based BNF trials (see Table S2). These data were chosen as our target variables used for model evaluation. In addition, to convert plant C mass
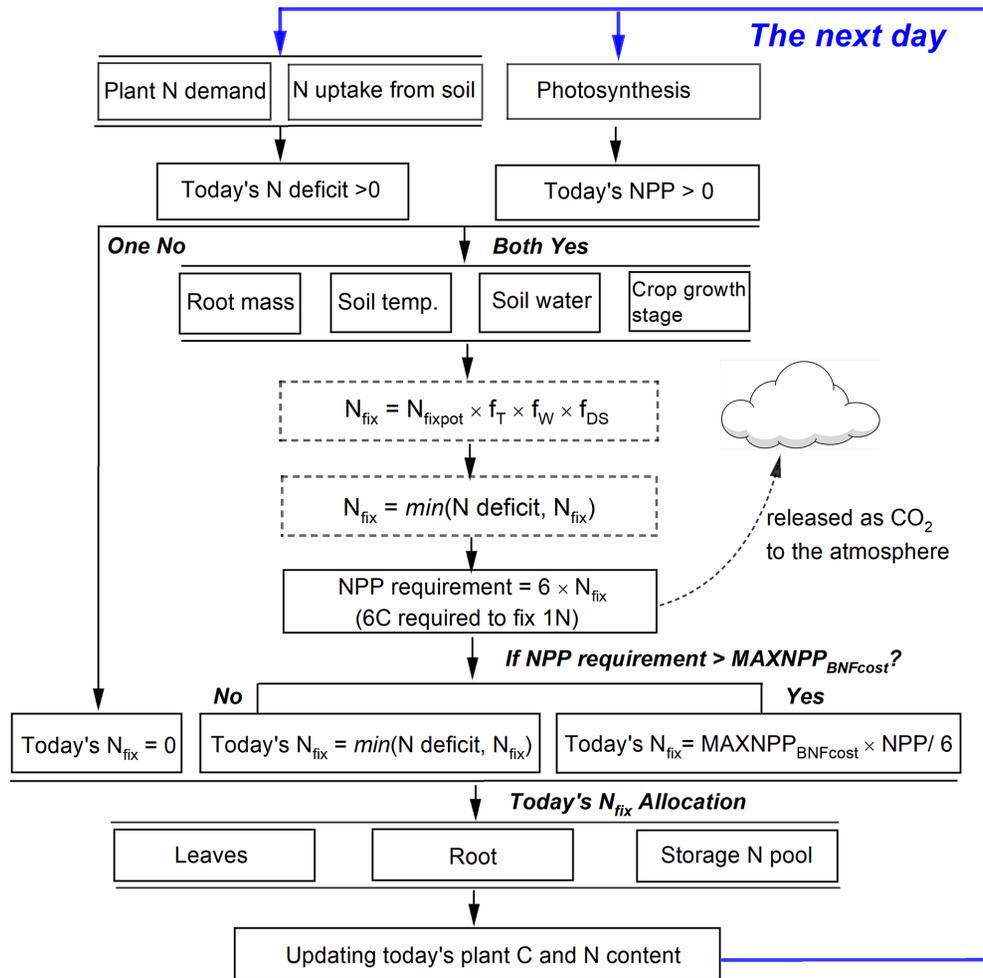
Figure 2. Representation of the N fixation route used in grain legumes in LPJ-GUESS. Today's N deficit is calculated as the difference between plant N demand and soil mineral N uptake. N_{fix} in dotted boxes represents intermediate values.

to dry matter, a conversion factor of 2.0 was used (Smith et al., 2014). Dry weight was converted to wet weight by assuming a water fraction of 0.13 in the grain legumes (Córdova et al., 2019).

Since specific leaf area (SLA) and target grain C : N ratio play a very important role in determining N uptake and N retranslocation to grain during seed-filling in the model (Olin et al., 2015), we implemented two simulations to explicitly explore model performance across all sites. For "site-specific" simulations, the reported SLA and grain C : N ratio listed in Table S2 were adopted for the simulation (for sites for which these were available). For "global uniform" parameter simulations, SLA was set to 40 and 45 $\text{m}^2\,\text{kg}^{-1}$ C (Penning de Vries et al., 1989), and the target grain C : N ratio was represented as a constant of 8 for soybean and 10 for faba bean (Kattge et al., 2020). These values were also used for global-scale simulations.

Due to the unavailable information on weather data at the majority of the sites evaluated, gridded daily climate data for

air temperatures (maximum, minimum, and mean), precipitation, and solar radiation were used from GSWP3-W5E5 (Dirmeyer et al., 2006; Lange, 2019; Cucchi et al., 2020), chosen for the $0.5° \times 0.5°$ grid cell representative for each experimental site. We compared model-required input variables from GSWP3-W5E5 with observations at three sites, finding that the gridded climate data had fairly good agreement with weather records in the field, despite some solar radiation deviations between two datasets for individual days over the experimental period (Fig. S2). There was no information on land-use and management practices in years preceding the experiments at most sites. Therefore, to maintain soil N and C pools in equilibrium after model spin-up, we decided to implement a common cropping system of maize–legume rotation annually from 1901 to the year before the trial start, with no N fertilizer applied to legumes. Over the trial period, the management practices were implemented according to information provided in the literature (Table S2). In addition, site-specific soil physical properties, such as frac-

Table 1. Overview of BNF-related variables and parameters used in the model for soybean and faba bean.

Parameter	Description	Soybean	Faba bean	Unit	Reference
N deficit	plant N demand minus soil N uptake	dynamic	dynamic	$\mathrm{g\,N\,m^{-2}\,d^{-1}}$	
NPP	net primary productivity	dynamic	dynamic	$\mathrm{g\,C\,m^{-2}\,d^{-1}}$	
$N_{\mathrm{maxfixpot}}$	maximum nitrogen fixation rate of roots	0.03	0.03	$\mathrm{g\,N\,g^{-1}}$ root DM	Soussana et al. (2002), Eckersten et al. (2006), Boote et al. (2008)
$\mathrm{DM_{root}}$	root dry matter	dynamic	dynamic	$\mathrm{g\,root\,DM\,m^{-2}}$	
C cost	carbon cost per unit fixed N	6	6	$\mathrm{g\,C\,g^{-1}\,N}$ fixed	Ryle et al. (1979), Boote et al. (2008), Kaschuk et al. (2009)
T	soil temperature at a depth of 25 cm	dynamic	dynamic	°C	
T_{min}	the minimum temperature for the start of N fixation	5	1	°C	Boote et al. (2008)
T_{optL}	lower bound of optimal temperature for N fixation	20	16	°C	Boote et al. (2008)
T_{optH}	upper bound of optimal temperature for N fixation	35	25	°C	Boote et al. (2008)
T_{max}	the maximum temperature for the stop of N fixation	44	40	°C	Boote et al. (2008)
W_f	relative soil water content in the top layer (0–50 cm)	dynamic	dynamic	–	
W_a	lower bound of water content below which N fixation is fully limited by soil water deficit	0.2	0	–	Robertson et al. (2002)
W_b	upper bound of water content above which N fixation is not inhibited by water content	0.8	0.5	–	Robertson et al. (2002)
φ_1	coefficient of soil water content	−0.33	0	–	Robertson et al. (2002)
φ_2	coefficient of soil water content	1.67	2	–	Robertson et al. (2002)
NDS	normalized crop development stage	dynamic	dynamic	–	Wang and Engel (1998)
$\mathrm{NDS_{min}}$	the minimum development stage for the start of N fixation	0.1	0.1	–	Bouniols et al. (1991)
$\mathrm{NDS_{optL}}$	lower bound of development stage for N fixation	0.3	0.3	–	Bouniols et al. (1991)
$\mathrm{NDS_{optH}}$	upper bound of development stage for N fixation	0.7	0.6	–	Bouniols et al. (1991)
$\mathrm{NDS_{max}}$	the maximum development stage for the stop of N fixation	0.9	0.8	–	Bouniols et al. (1991)

tions of sand, silt, and clay, were also used as forcing to further compute corresponding soil water characteristics in the model (Olin et al., 2015).

2.4.2 Global yields and BNF rate

To evaluate the model's ability to simulate legume yields and BNF on a large scale, national crop yield statistics from FAOSTAT (http://www.fao.org/faostat/en/#data/QC, last access: 9 May 2021) were collected and compared with modeled output. Furthermore, Peoples et al. (2009) divided N fixation data for widely grown legume crops collated from a range of published sources into different geographical regions. In order to compare our simulated BNF with the literature-based records, each simulated $0.5° \times 0.5°$ grid cell was classified to be in one of the 10 regions given in Table 1 in Peoples et al. (2009) (Fig. S3).

For regional comparison, the modeled gridded yield and BNF rate were aggregated to national and continental scales, respectively, using information on crop-specific cover area in the spatial pattern (described below):

$$\mathrm{Var_{region}}$$
$$= \frac{\sum_{i=1}^{n}\left[(\mathrm{Var_{rain}})_i \times (\mathrm{Area_{rain}})_i + (\mathrm{Var_{irri}})_i \times (\mathrm{Area_{irri}})_i\right]}{\sum_{i=1}^{n}\left[(\mathrm{Area_{rain}})_i + (\mathrm{Area_{irri}})_i\right]}, \quad (15)$$

where Var is yield or BNF rate; $\mathrm{Var_{region}}$ is the aggregated result in a given region; i is the grid cell number in that region, ranging from 1 to n; $\mathrm{Var_{rain}}$ and $\mathrm{Var_{irri}}$ represent the modeled yield or BNF rate under rain-fed and irrigated conditions, respectively; and $\mathrm{Area_{rain}}$ and $\mathrm{Area_{irri}}$ are the crop-specific rain-fed and irrigated areas used in simulations, respectively.

As land-use and land cover input, data from LUH2 (Land-Use Harmonization 2, Hurtt et al., 2020) with fractions of

cropland, pasture, and natural vegetation at each grid cell were adopted, spanning from 1901 to 2014 in 0.5° resolution. The fractional cover of different crop species was derived from MIRCA (Monthly Irrigated and Rain-fed Crop Areas; Portmann et al., 2010). Since no detailed information was available on the growth distribution of the faba bean, the "pulse" fraction in MIRCA was used as input instead, and pulse country-level yield statistics provided by FAOSTAT (2021) were collected to compare with faba bean outputs by LPJ-GUESS. As information on cropland soil characteristics, data in the top layer (30 cm) were derived from the GGCMI (Global Gridded Crop Model Intercomparison) phase 3 soil input dataset (Volkholz and Müller, 2020). In general, although the total cropland cover in a grid cell could change annually over the course of the simulation, the relative fractions of each crop species within that cover fraction were held constant.

In terms of timing of N fertilizer application, a recent meta-analysis conducted by Mourtzinis et al. (2018) indicated that splitting N application between planting and the early reproductive stage resulted in significantly greater soybean yields than a single application. Mineral N fertilizer for legumes in the model was thus split into two equal applications at the time of sowing (DS = 0) and flowering (DS = 1.0). Manure was added to soils at the time of sowing as a single application to reflect real-world practices that account for the time required for manure N to be made available to plants. Data sources for mineral N fertilizer and manure over the period 1901–2014 were derived from Ag-GRID (AgMIP GRIDded Crop Modeling Initiative; Elliott et al., 2015 and Zhang et al., 2017, respectively) (Fig. S4).

2.5 Statistical methods

In order to quantify the agreement between modeled and observed variables, the coefficient of determination (adjusted R^2), relative bias (RB, Eq. 16), absolute bias (AB, Eq. 17), and the root mean square error (RMSE, Eq. 18) were computed:
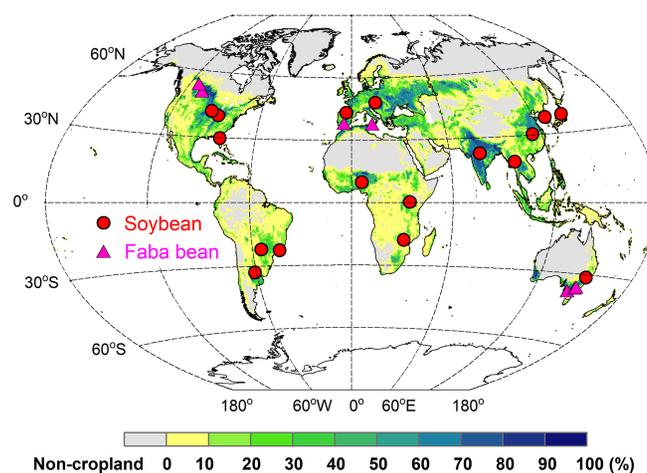
$$RB = \frac{M_i - O_i}{O_i} \times 100\%, \tag{16}$$

$$AB = \frac{|M_i - O_i|}{O_i} \times 100\%, \tag{17}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(M_i - O_i)^2}, \tag{18}$$

where M_i and O_i indicate modeled and observed values, and n is the number of observations. To evaluate the fit of the interannual variability of modeled and reported yields on the country level, the standard deviation (SD) and Pearson correlation coefficient (r, Eq. 19) were calculated:

$$r = \frac{\sum_{i=1}^{n}(M_i - \overline{M})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{n}(M_i - \overline{M})^2 \sum_{i=1}^{n}(O_i - \overline{O})^2}}, \tag{19}$$

Figure 3. Spatial distribution of soybean (red circles) and faba bean (magenta triangles) sites used for BNF evaluation. The map background is cropland fraction (%) averaged over 1996–2005 at the resolution of 0.5° × 0.5°, derived from the LUH2 dataset (Hurtt et al., 2020).

where $\overline{M}$ and $\overline{O}$ represent modeled and observed mean, and n is the number of reported years.

3 Results

3.1 Model evaluation at site scale

3.1.1 Model performance across all sites

In order to examine model performance in simulating BNF-related variables across all grain legume sites described in Table S2, we compiled six widely measured variables related to N fixation at harvest, as shown in Fig. 4. Modeled yields generally agreed well with observations, especially in the site-specific simulation setup. These had a higher regression slope (0.83) and lower absolute bias (28 %) compared with the global uniform simulation setup (Fig. 4a). N content in grains and shoots showed lower agreement, with simulated values underestimating the observations for most sites (Fig. 4b–c), likely arising from two important N sources to grain legumes not being captured well by the model (i.e., soil N uptake and BNF, shown in Fig. 4d–e). The global uniform run did not capture observed N fixation well, with a regression slope of 0.22 and absolute bias of 39 %. The simulated BNF compared to observations was notably improved when using site-specific parameters, with the regression slope increasing to 0.41 and the absolute bias being reduced to 31 % (Fig. 4e). The field-based measurements showed that the N derived from the atmosphere (%Ndfa) was the main contributor to the legumes' total N uptake, ranging from 15 % to 95 %, with a mean of 64 % across all field trials. LPJ-GUESS generally captured the mean response well, with simulated %Ndfa being 60 % and 58 % in the site-specific and global

uniform runs, respectively, despite several extreme disagreements at several faba bean sites (Fig. 4f).

A linear relationship between legume yields and the rate of BNF was found across a range of field sites in this study (Fig. S5a). Simulations from LPJ-GUESS mostly captured the close correlation between these variables, with R^2 ranging 0.46–0.63 ($p < 0.001$) in both runs, which is not far from the measured value of 0.67 (Fig. S5a). Linear regression parameters (i.e., slope and intercept) in both runs were close to the observations, indicating that the model reproduces the N fixation effect on yield well for individual sites.

A negative exponential relationship was observed between N fertilizer application rate and N fixation across the field trials (Fig. S5b). LPJ-GUESS reasonably reproduced the decreased trend of BNF to N fertilizer increase, with similar fitting functions to observations, although higher N fixation rates were modeled in the highest-fertilized trial (600 kg N ha^{-1}) compared with measurements (Fig. S5b).

3.1.2 Response to irrigation

The ability of the model to simulate the observed response of soybean tissue biomass and N mass to irrigation management was examined using data from an experiment with rain-fed and irrigated treatments in Florida, USA (82.4° W, 29.6° N; see Table S2). Since the timing and quantity of irrigation were not reported in the literature (DeVries et al., 1989a, b), we assumed that soybean was irrigated automatically when it experienced water stress in the model, with the amount of plant water deficit as supplemental irrigation.

The mean observed grain yields at harvest were 2.0 and 2.9 t ha^{-1} under rain-fed and irrigated conditions, respectively, whereas the modeled yields were 1.9 and 2.5 t ha^{-1} for the site-specific parameter run and 1.6 and 2.1 t ha^{-1} for the global uniform parameter run, suggesting good model performance for rain-fed crops but an underestimation of the effect of irrigation on yields (Fig. 5a). Grain dry matter over the cropping season was simulated to increase by 32 % and 45 % on average in response to irrigation in the site-specific and global uniform runs, respectively. The observations show a similar response but with a higher increase of 75 %. The modeled increase in grain N content caused by irrigation also showed good agreement, with an increase of 35 %–58 % in both runs, in line with the observed response of 42 % (Fig. 5b).

The model generally reproduced observed leaf biomass and N mass better than the total aboveground production under rain-fed and irrigated treatments, with higher accuracy in the site-specific run. Over the growing season there was an obvious underestimation of the total aboveground production of biomass for both runs (Fig. 5a). This may be partially due to the fact that LPJ-GUESS at this point does not model soybean hulls, which account for $\sim$ 15 %–20 % of the total aboveground dry matter at harvest in the US soybean rain-fed cropping system (Córdova et al., 2020). The ob-

served increase in shoot and leaf biomass due to water supply was 19 % and 21 %, respectively. In comparison, the site-specific parameterized model resulted in increases of 13 % and 14 %, respectively (15 % and 14 % for the global uniform parameter run, see Fig. 5b). Overall, the observed soybean tissue biomass and N content under rain-fed and irrigated conditions, as well as their response to irrigation management, were captured reasonably well by the model at the US Florida site, despite some deviations from observations in some cases.

3.1.3 Response to nodulating soybean

In Zapata et al. (1987), two field trials with non-nodulating and nodulating soybean were conducted in Seibersdorf, Austria (16.5° E, 48.0° N; see Table S2), resulting in different plant C and N production at various growth stages. As described in Sect. 2.3, the nodulation process of legumes has not yet been implemented in LPJ-GUESS; we thus switched off (on) the BNF function in the model to simply represent the non-nodulating (nodulating) soybean experiment.

During the growing season, yield and grain N mass in the field trials increased rapidly after the vegetative stage, peaking around harvest. Simulations from LPJ-GUESS mostly captured those seasonal dynamics and the response to nodulating soybean (Fig. 6a–b): the modeled increase in yield and grain N mass due to nodulation was 34 % and 51 % in the site-specific run (34 % and 45 % in the global uniform run), respectively, in line with the observed response of 20 % and 41 % at harvest (Table 2), which suggests appropriate sensitivity of yield and N content in grain to N addition from N fixation. Similarly, the model generally reproduced the observed seasonal pattern of shoot N mass well, but with some underestimations in the nodulation trial (Fig. 6c).

Accumulated soil N uptake was captured reasonably well over the entire growing season, with higher accuracy at harvest in the global uniform simulation (Fig. 6d). Measured mineral N uptake from soils declined on average by 25 % in response to nodulation. In comparison, the simulated reduction in uptake was 50 % and 46 % for the site-specific and global uniform runs (Table 2). The BNF rates were low at the early growth stages when nodules were still establishing and increased rapidly between floral initiation and the early seed-filling, after which nodule senescence occurred and the increase in N fixation rate declined until physiological maturity (Fig. 6e). Simulations from LPJ-GUESS reproduced the seasonal pattern of N fixation with some overestimations in the accumulated BNF at the end of the growth period; the site-specific and global uniform runs simulated 113 and 140 kg N ha^{-1}, respectively, compared to the measured value of 103 kg N ha^{-1} (Table 2).
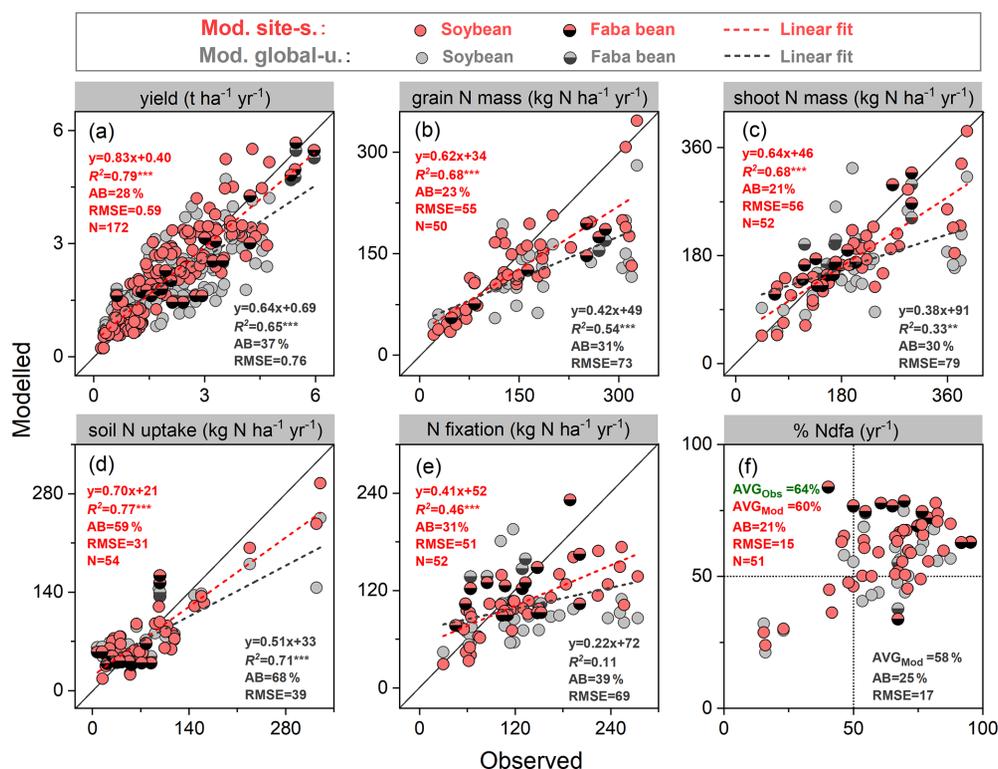
Figure 4. Comparison of modeled and observed yield **(a)**, grain N mass **(b)**, shoot N mass **(c)**, soil N uptake **(d)**, BNF **(e)**, and %Ndfa (the proportion of plant N derived from the atmosphere) **(f)** at harvest across all soybean and faba bean sites. Filled red and grey circles depict the "site-specific" and "global uniform" runs, respectively. The dashed line is a fitted linear regression with red for site-specific and grey for global uniform; *** and ** denote regressions statistically significant at $p = 0.001$ and 0.01, respectively; AB is absolute bias (Eq. 17), represented in percent (%); the unit of RMSE is the same as the associated variable; AVG in **(f)** is the averaged value of %Ndfa across all field trials.
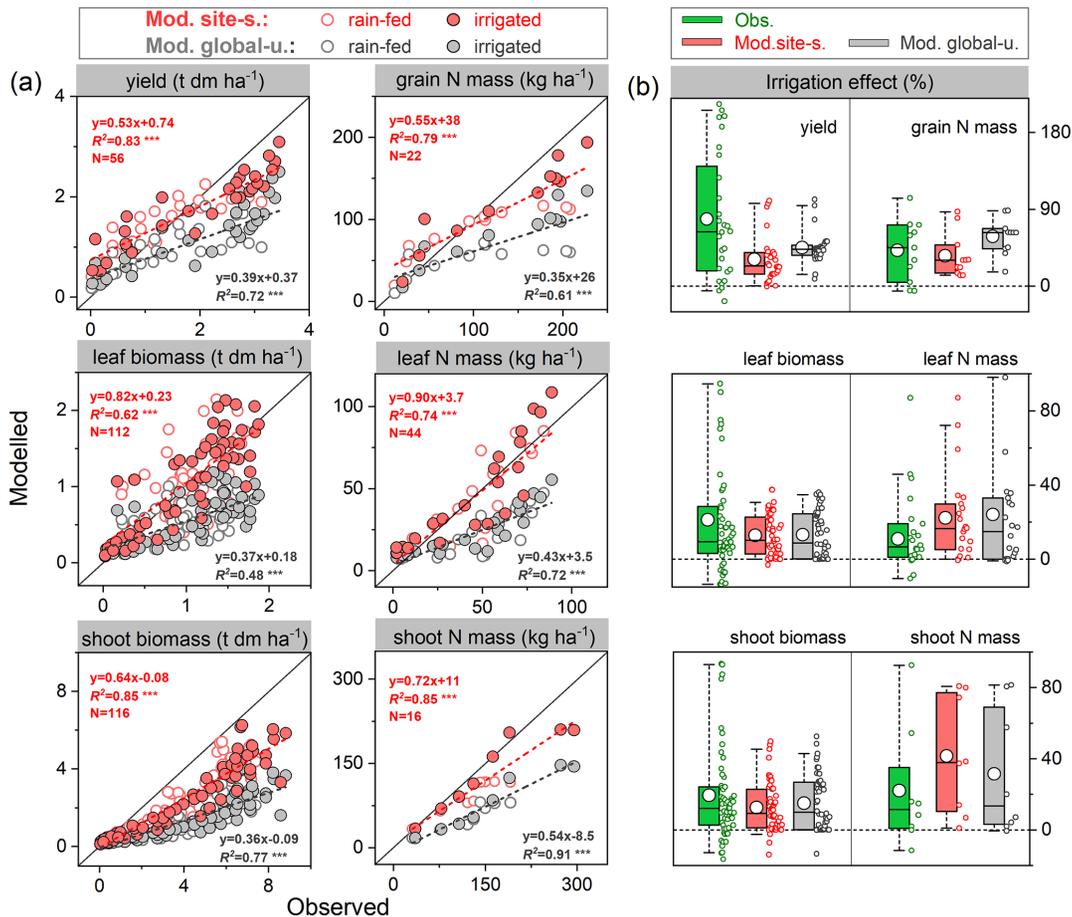
Table 2. Comparison of modeled and observed yield ($t\,ha^{-1}$), grain N mass ($kg\,N\,ha^{-1}$), shoot N mass ($kg\,N\,ha^{-1}$), soil N uptake ($kg\,N\,ha^{-1}$), and N fixation rate ($kg\,N\,ha^{-1}$) from a soybean nodulation and non-nodulation experiment at harvest. The observed data were compiled using Tables 2–4 in Zapata et al. (1987).

	Nodulation			Non-nodulation			Nodulation effect (%)		
	Obs.	Mod. site-s.	Mod. global-u.	Obs.	Mod. site-s.	Mod. global-u.	Obs.	Mod. site-s.	Mod. global-u.
Yield	3.01	3.24	3.06	2.42	2.41	2.29	20	34	34
Grain N mass	162	166	148	115	110	102	41	51	45
Shoot N mass	222	198	181	158	134	138	41	48	31
Soil N uptake	119	76	86	158	152	159	−25	−50	−46
N fixation	103	140	113	–	–	–	–	–	–

3.1.4 Response to N fertilizer in faba bean

In the N fertilizer experiment from Mínguez et al. (1993), four field trials were compared with N applications between 0 and $300\,kg\,N\,ha^{-1}$ at three crop growth stages and two faba bean varieties grown in a Mediterranean climate (Spain, 4.8° W, 37.9° N; see Table S2). Over the entire growing season, leaf biomass and N content in the field trials increased until around May, after which leaf senescence started and biomass and N content declined (Fig. 7a–b). The model broadly reproduced these seasonal patterns and the response to different N application rates. The largest difference between modeled and measured leaf biomass was found at the end of the growing season as a result of the simulated leaf senescence rate being much lower than derived from measurements (Fig. 7a). In addition, the simulations showed modeled leaf N mass to decline rapidly during the late reproductive phase. This can be attributed to the transfer of N from vegetative parts to grain because of the high N demand in seeds during the grain-filling period.

Figure 5. Comparison of modeled and observed soybean tissue biomass and N mass **(a)** and their responses to irrigation management **(b)** compared with those grown at rain-fed conditions. Red and grey circles depict "site-specific" and "global uniform" run, respectively; the dashed line is fitted linear regression; *** denotes the regression statistically significant at $p = 0.001$. Box plots in **(b)** denote the 5th and 95th percentiles with whiskers, median and interquartile range with box lines, and mean with a white dot (all data distributed next to the box). The seasonal data at each phenological stage for tissue biomass are available from 1978–1979 and 1984–1985 with rain-fed and irrigated treatments; those for N mass are available for 1979 and 1984, while the seasonal shoot N mass is only available for 1984.

As seen in Fig. 7c, modeled soil N uptake was stimulated by soil mineral N availability, with an increase of 120 %–160 % compared to the unfertilized treatment. In contrast, fixing N from the atmosphere was constrained in the presence of elevated levels of soil mineral N, with a reduction of 15 %–20 %. The total N uptake for the cropping season 1987–1988 was observed to only increase by 3 % in response to N application as a consequence of the inoculation implemented in the unfertilized treatment (Mínguez et al., 1993). By contrast, LPJ-GUESS produced relatively large increases of 14 %–16 % in both runs, resulting in the observed increase in plant biomass and N mass accumulation caused by N addition being largely overestimated in the model (Fig. 7c).

3.2 Model evaluation at global scale

3.2.1 Attained yields

Using the global uniform parameters described in Sect. 2.4.1, combined with the time-dependent gridded N fertilizer dataset introduced in Sect. 2.4.2, we simulated soybean and all pulses (applying the faba bean parameterization, see Sect. 2.4.2) at a global scale. We computed data for the period 1996–2005, since crop-specific fractional cover from the MIRCA dataset was available for the year 2000 (Portmann et al., 2010).

Modeled yields in the top 10 soybean-producing countries showed good agreement, with a higher R^2 of 0.52 ($p < 0.001$) and lower RMSE value of 0.8 t ha^{-1} yr^{-1} when low-productivity countries (defined as all countries not belonging to the top 10 producer countries) were excluded. With all producer countries included, R^2 of 0.17 ($p < 0.001$)
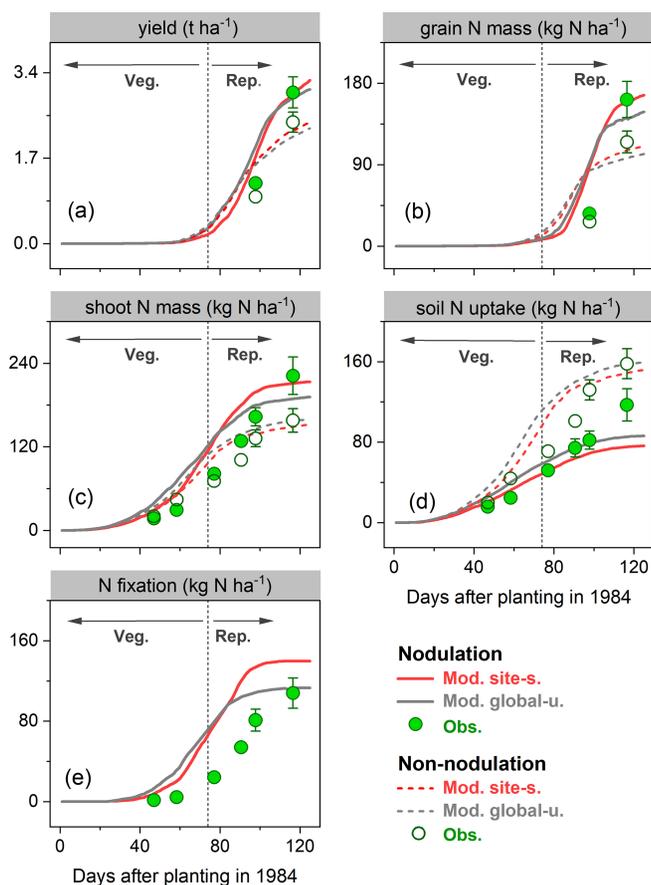
Figure 6. Observed (circles) and modeled (lines) yield **(a)**, grain N mass **(b)**, shoot N mass **(c)**, soil N uptake **(d)**, and BNF **(e)** for a field site in Austria (Zapata et al., 1987) for the cropping season 1984 with nodulating and non-nodulating soybean. The observed values of soil N uptake and BNF across all growth stages were calculated based on Fig. 1 given in Zapata et al. (1987), and the vertical bars represent the standard error of a four-replicate mean in the original literature. Veg. and Rep. indicate vegetative and reproductive growth phase, respectively.

and RMSE of $1.4\,t\,ha^{-1}\,yr^{-1}$ were found (Fig. 8a). LPJ-GUESS generally tended to overestimate the reported yield for most countries where soybean production is low (e.g., most African countries, see Fig. 9a), with a mean relative bias in such countries of 81 % (Fig. 8a). Modeled low yields were found in some arid and semi-arid countries (e.g., Egypt, Iran, and Turkey), with the underestimation spanning from 10 %–70 % (Fig. 9a). Overestimated yields were also found when comparing simulated yields using the faba bean parameterization against FAO-reported values for pulses in general, with an overestimation also visible for some of the top producing countries (Fig. 8b). Likely, the higher yields simulated by LPJ-GUESS arise from the fairly high N fixation capacity simulated with the faba bean parameterization (see Sect. 3.2.2) and the wide distribution of pulses worldwide, which grow under a broad range of climate and soil conditions.

A good fit of the interannual variability of modeled and reported yields is a further indicator of model performance. Despite the deviation between the model and observations for individual years, simulated variation in soybean yield over the period 1981–2016 matched reported yields well among the top 10 producer countries – especially in Argentina, India, and China – with a high Pearson correlation coefficient (r) around 0.60 ($p < 0.001$) and similar standard deviations (Fig. 9). The degree of yield variability between years was larger than seen in the FAO records, especially in the US, Canada, and Italy (Fig. 9), indicating high sensitivity of modeled soybean yield to changing environmental factors on spatial scales, such as weather, N fertilizer application rates, and climate-related N fixation.

3.2.2 N fixation and %Ndfa

The modeled spatial pattern of soybean N fixation showed large spatial variation (Fig. 10a). Modeled BNF rates as high as $250\,kg\,N\,ha^{-1}\,yr^{-1}$ were found in western South America and most of Africa, where neither water nor temperature was a critical limitation for N fixation. Moreover, the relatively low fertilizer application in Africa (0–$20\,kg\,N\,ha^{-1}\,yr^{-1}$, Fig. S4b) leaves a nitrogen deficit that causes enhanced soybean N fixation. In contrast, in arid and semi-arid regions, soil water constrains BNF, while temperature limitation is seen in high latitudes and alpine areas (e.g., Andes in Peru). BNF rates in most regions (South Asia, West Asia, sub-Saharan Africa, and northwestern China) were as low as $50\,kg\,N\,ha^{-1}\,yr^{-1}$, particularly in Pakistan and northern India, where simulated BNF is severely constrained by the extreme high temperature over the cropping season. The eastern United States, Europe, southern China, and central-western Brazil showed intermediate fixation rates, which were greater than $150\,kg\,N\,ha^{-1}\,yr^{-1}$. Overall, the spatial variation of the modeled legume BNF rate reflects to a large degree the spatial climate patterns, in addition to N fertilizer application. The low modeled %Ndfa of 45 ± 3 % in East Asia may reflect high N uptake from soils in response to substantial fertilizer investment in China (80–$180\,kg\,N\,ha^{-1}\,yr^{-1}$, Fig. S4b) over the past 40 years. In contrast, the modeled %Ndfa in Africa – with lower N application rates – was as high as 70 ± 3 %, although still lower than the reported mean value of 77 % (Table 3). The spatial response of N fixation rate to climate constraining factors (i.e., soil temperature and water) is shown for pulses in Fig. S6.

At a regional scale, the modeled outputs compare well with N fixation rates from the literature (Fig. 10b–f, Table 3). For example, in South America and North America, both major soybean-producing regions, simulated BNF rates were 156 ± 14 and $127\pm44\,kg\,N\,ha^{-1}\,yr^{-1}$ over the period 1981–2016, respectively, compared with literature-derived values of 136 and $144\,kg\,N\,ha^{-1}\,yr^{-1}$ (Peoples et al., 2009).
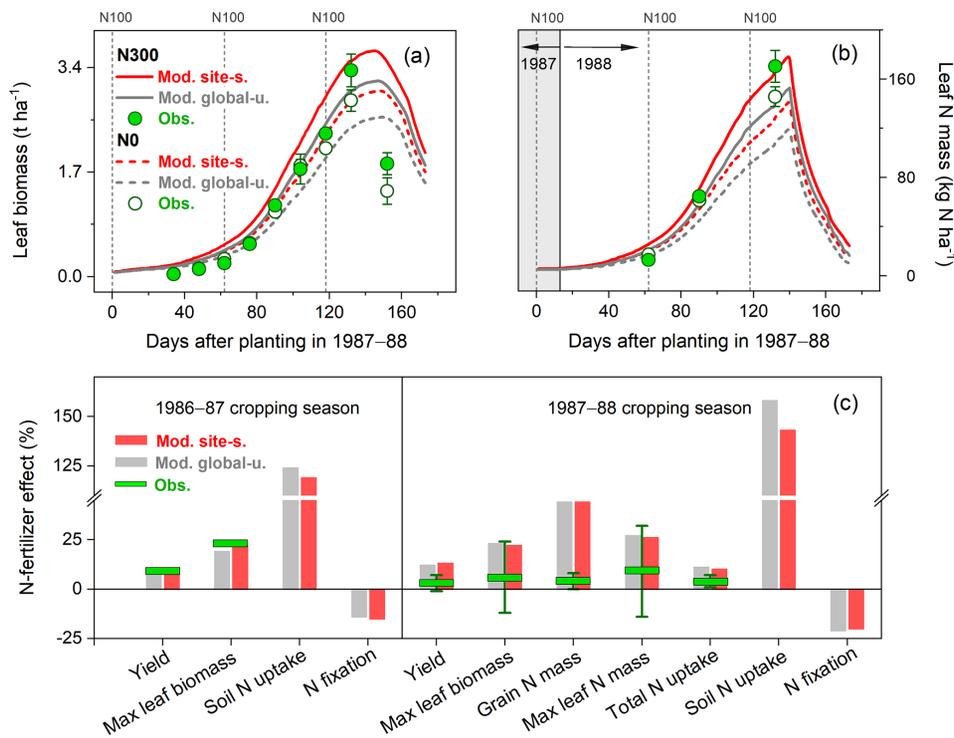
Figure 7. Observed and modeled seasonal pattern of leaf biomass **(a)** and leaf N **(b)** of faba bean in Spain for the cropping season 1987–1988, with two different levels of N fertilizer input (0 and 300 kg N ha^{-1} represented as N0 and N300, respectively), and the response of faba bean yields and N uptake to fertilized treatment at harvest **(c)** compared with those grown in unfertilized conditions in the 1986–1987 and 1987–1988 cropping seasons. The observed values were derived from the average of two faba bean varieties described in Mínguez et al. (1993), and their measured ranges are shown by the vertical bars. The vertical dashed lines in **(a)**–**(b)** represent the timing and amount of fertilizer applied in the N300 treatment.
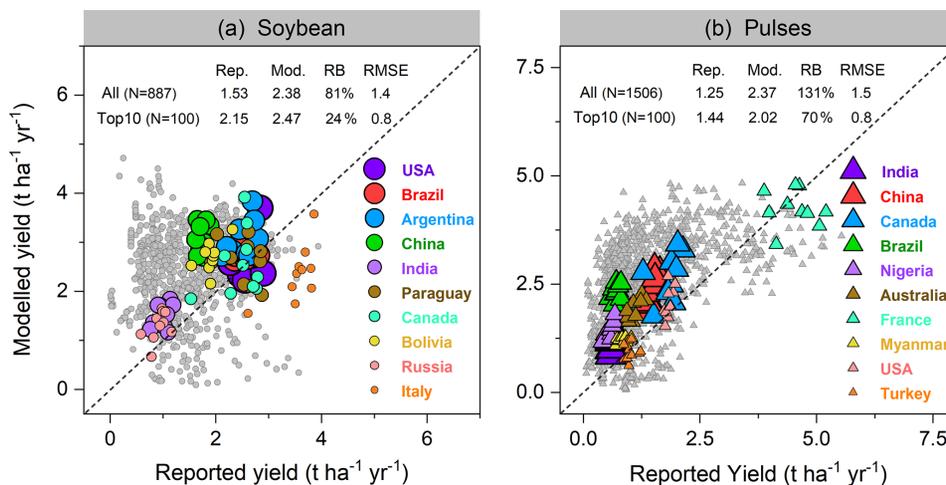
Figure 8. Per-country and per-year comparison of modeled yields of soybean **(a)** and pulses **(b)** against reported FAO statistics from 1996–2005. Each filled circle in **(a)** represents 1 year and one country; thus, a country can have up to 10 circles over 1996–2005. In total, 887 and 1506 country–year yield data points were used for comparison in soybean and pulses, respectively. The top 10 producer countries (shown in color) were chosen based on their total production over the same period, and marker size from large to small indicates their total relative production in descending order. Rep. and Mod. respectively denote reported and modeled yield (t ha^{-1} yr^{-1}) averaged from 1996–2005. RB is relative bias (Eq. 16), represented in percent (%). The unit of RMSE is the same as yield (t ha^{-1} yr^{-1}).
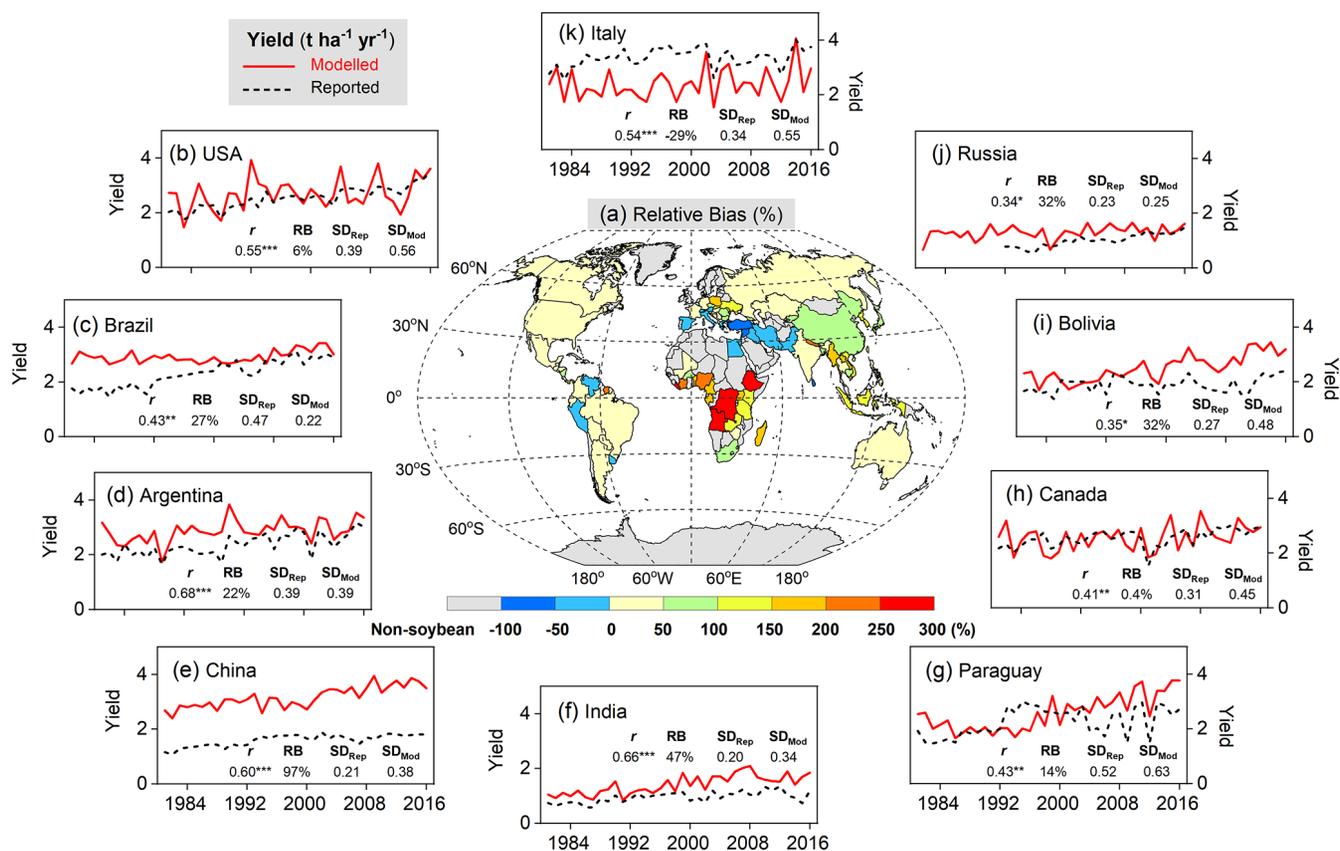
Figure 9. Comparison of simulated and FAO-reported yields on the country level averaged over 1996–2005 (**a**), as well as time series of modeled soybean yield (red solid line) and reported FAO statistics (black dashed line) in the top 10 producer countries over the period 1981–2016. The top 10 producer countries (**b–k**, in descending order) were chosen based on their total production from 1996–2005. r is the Pearson correlation coefficient (Eq. 19), where ***, **, and * denote the correlation as statistically significant at the $p = 0.001$, 0.01, and 0.05 level, respectively. RB is relative bias (Eq. 16), represented in percent (%). SD_{Rep} and SD_{Mod} denote respectively reported and modeled yield standard deviations ($t\,ha^{-1}\,yr^{-1}$) from 1981–2016.

Globally, the modeled soybean N fixation rate of $132 \pm 21\,kg\,N\,ha^{-1}\,yr^{-1}$ was reasonably consistent with the meta-analysis result of $111–125\,kg\,N\,ha^{-1}\,yr^{-1}$ in Salvagiotti et al. (2008) and the FAO-based estimate of $176\,kg\,N\,ha^{-1}\,yr^{-1}$ from Herridge et al. (2008). The contribution of N fixation to total N uptake in soybean was somewhat underestimated in several regions. A similar trend to underestimate reported %Ndfa was also found for pulses (Table 3).

Having large soybean-planting areas and high yields, South America and North America contributed 80 % of simulated global soybean N fixation, followed by East Asia, South Asia, and Europe (Table 3). Globally, simulated annual N fixed over the period 1981–2016 was $11.6 \pm 2.2\,Tg$ in soybean, which showed good agreement with the estimate of $16.4\,Tg\,N$ reported by Herridge et al. (2008) and the extrapolated result of $10.4\,Tg\,N$ estimated by Gelfand and Robertson (2015) based on US field trials. However, we modeled pulses to fix $5.6 \pm 1.0\,Tg\,N$ annually, which is almost 2 times higher than the $2.95\,Tg\,N$ estimated by Herridge et al. (2008). The difference in the case of pulses is most likely due to the low N

fixation rate used by Herridge et al. (2008) ranging from 23–$107\,kg\,N\,ha^{-1}\,yr^{-1}$, which is lower than the mean value of $119 \pm 15\,kg\,N\,ha^{-1}\,yr^{-1}$ modeled by LPJ-GUESS (Table 3).

4 Discussion

4.1 Model performance at site scale

The overall model agreement with measured legume yield and grain N mass was good across a range of field sites (Fig. 4). Values at harvest were on average about 20–30 % lower than values reported in the measurements (Fig. 4a–b). A similar small underestimation was found in the shoot N mass (Fig. 4c), indicating that the productivity is generally somewhat too low in the model. One factor contributing to the underestimation is that LPJ-GUESS applies a conversion factor of 2.0 from plant C mass to dry matter (Smith et al., 2014), which is $\sim 10\,\%$ lower than a published measurement of 2.24 reported in Osaki (1993). In addition, we found that the model underestimated aboveground biomass while
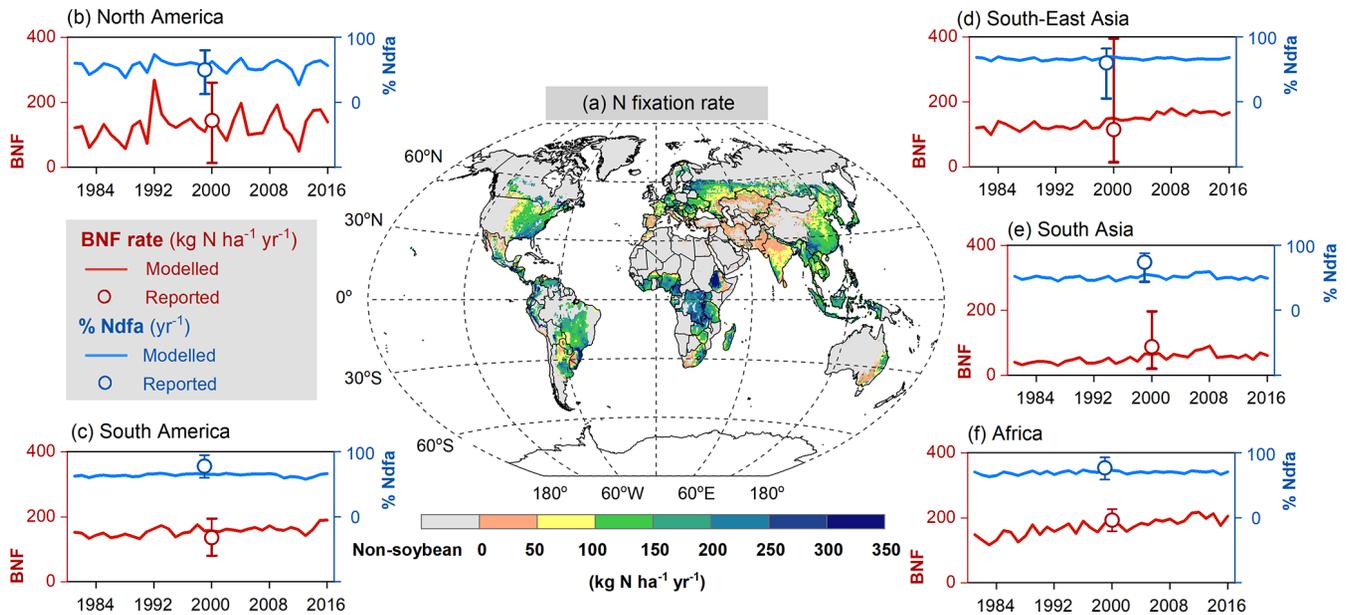
Figure 10. Map of soybean N fixation modeled by LPJ-GUESS averaged over 1996–2005 **(a)** and the comparison of simulated BNF rate (red line) and %Ndfa (blue line) with literature-reviewed data (open circle; Peoples et al., 2009) on a regional level **(b–f)**. Reported data shown in open circles do not represent specific years but the potential over time in Peoples et al. (2009); the vertical bars denote the range of estimations based on the original literature given in Table 1 in Peoples et al. (2009).

simultaneously overestimating belowground productivity at the three sites where measured root biomass was available. This could be addressed by adjusting the root : shoot allocation (i.e., modifying the daily assimilate partitioning function in grain legumes; Eq. 5), but this is currently prevented by the lack of sufficient observed root biomass information.

Modeled soil N uptake was sensitive to soil mineral N concentration and hence driven by fertilizer application rates (Figs. 7c, S5c). Generally, LPJ-GUESS tended to overestimate soil N uptake in regions where legumes were not fertilized or only lightly fertilized (Fig. S5c). This might be partially due to the selected legume cultivars at the experimental plots, which have been reported to have low mineral N uptake potential (Gan et al., 2002, 2003; Santachiara et al., 2017, 2018). Moreover, the saturation effect of mineral N concentration on N uptake implemented in the model might result in the discontinuation of N uptake when soil-available N is abundant (Zaehle and Friend, 2010; Wårlind et al., 2014). Under high fertilization rates (up to 260–600 kg N ha^{-1}, Fig. S5c), a strong underestimation in soil N uptake was expected because of the modeled saturation response to high soil mineral N, resulting in little change in the level of soil N uptake no matter how much N fertilizer was applied.

Adding mineral N to the soil in LPJ-GUESS can increase soil N uptake, reducing the plant's N deficit and therefore also reducing the upper limit of the daily N fixation rate (Fig. 2). Although the modeled negative relationship between fertilizer application rates and N fixation showed generally

good agreement with the observed response across a range of field sites, the simulated BNF rates in the highly fertilized trials (i.e., 260–600 kg N ha^{-1}, Fig. S5b) were about 50 %–80 % higher than the measured values (Figs. 4e, S5b). This might be partially explained by the underestimation in soil N uptake under higher N concentration, resulting in plant N demand remaining very high and substantial N still being fixed. The large discrepancies between modeled and observed N uptake in the highly fertilized treatments suggest that the N uptake representation in LPJ-GUESS should be further improved. A step forward could be to incorporate the inhibitory effects of soil mineral N content on N fixation into the model (Chen et al., 2016; Wu et al., 2020), since experimental evidence indicates that high soil mineral N not only affects plant N uptake in roots, but also depresses legume nodule initiation, nodule size, and specific nodule activity, therefore reducing the amount of N fixation from the atmosphere (Herridge et al., 1984; Purcell and Sinclair, 1990; Thornley and Cannell, 2000).

The percentage of plant N derived from the atmosphere (i.e., %Ndfa) is a key parameter required for quantifying N fixation in the field and varies widely, caused by differences in climate, soil type, and degree of N fertilization (Herridge et al., 2008). LPJ-GUESS captured the range and mean value of %Ndfa well across different field trials, with some disagreements, especially for faba bean (Fig. 4f). An underestimated %Ndfa is likely caused by the combined effects of underestimated N fixation (Fig. 4e) and overestimated soil N uptake (Fig. 4d). Nevertheless, we found modeled %Ndfa to

Table 3. Modeled continent-level biological N fixation rate, the proportion of plant N derived from the atmosphere (%Ndfa), and total N fixation in soybean and pulses for the time period 1981–2016 compared to estimates from the literature, with the reported range in brackets. The modeled results are represented as mean ±1 standard deviation.

	Soybean					Pulses				
	N fixation rate (kg N ha⁻¹ yr⁻¹)		%Ndfa (yr⁻¹)		Total N fixation (Tg N yr⁻¹)	N fixation rate (kg N ha⁻¹ yr⁻¹)		%Ndfa (yr⁻¹)		Total N fixation (Tg N yr⁻¹)
	Reported	Modeled	Reported	Modeled	Modeled	Reported	Modeled	Reported	Modeled	Modeled
South Asia	88[a] (21–197)	53 ± 14	74[a] (44–88)	51 ± 3	0.4 ± 0.1	–	62 ± 12	–	52 ± 2	0.9 ± 0.2
Southeast Asia	115[a] (0–400)	141 ± 22	60[a] (0–82)	66 ± 2	0.2 ± 0.0	–	139 ± 16	–	69 ± 1	0.3 ± 0.1
Africa	193[a] (159–227)	172 ± 25	77[a] (65–89)	70 ± 3	0.2 ± 0.1	–	157 ± 21	–	70 ± 1	1.9 ± 0.5
North America	144[a] (14–311)	127 ± 44	50[a] (13–80)	56 ± 9	4.9 ± 1.7	118[b] (13–252)	137 ± 21	74[b] (60–92)	59 ± 3	0.6 ± 0.1
South America	136[a] (80–193)	156 ± 14	78[a] (60–95)	64 ± 2	4.5 ± 1.1	–	157 ± 18	–	66 ± 3	0.5 ± 0.1
East Asia	–	101 ± 16	–	45 ± 3	1.2 ± 0.2	–	114 ± 17	–	49 ± 4	0.4 ± 0.1
Central Asia	–	63 ± 19	–	36 ± 6	0.0 ± 0.0[e]	–	104 ± 21	–	59 ± 4	0.0 ± 0.0[e]
West Asia	–	27 ± 7	–	14 ± 3	0.0 ± 0.0[e]	100[b] (78–133)	65 ± 10	69[b] (63–76)	35 ± 4	0.2 ± 0.0
Europe	–	117 ± 17	–	54 ± 4	0.2 ± 0.1	153[b] (73–211)	177 ± 26	74[b] (60–92)	63 ± 4	0.6 ± 0.1
Oceania	–	78 ± 27	–	38 ± 9	0.0 ± 0.0[e]	143[b] (82–216)	126 ± 23	82[b] (69–89)	37 ± 6	0.4 ± 0.1
Global	111–176[a,c,d]	132 ± 21	52–68[a,c,d]	57 ± 4	11.6 ± 2.2	107–129[b,d]	119 ± 15	75[b,d]	60 ± 1	5.6 ± 1.0

[a] Soybean data in Peoples et al. (2009). [b] Faba bean data in Peoples et al. (2009). [c] Salvagiotti et al. (2008). [d] Herridge et al. (2008). [e] The values do not represent zero grain-legume-planting area in that region.

decline with increasing N fertilizer application, which is also the observed response in the field trials. A negative correlation between %Ndfa and fertilizer application rates was also reported by Salvagiotti et al. (2008). These results all suggest that LPJ-GUESS is able to effectively capture the observed overall patterns of soil mineral N uptake and N fixation in grain legumes and their responses.

Since the SLA and C : N ratio of plant organs play a vital role in determining N uptake when modeling vegetation C–N dynamics (Olin et al., 2015), it is to be expected that applying measured values for site-scale modeling resulted in much better agreement when comparing simulation results to measurements (Figs. 4–7). Remaining discrepancies between modeled and observed N-cycle variables may reflect missing processes in the model, such as inoculation effectiveness, phosphorus limitation, and soil acidity, especially in terms of inoculant application. Field experiments have shown that proper inoculation of rhizobia promotes nodulation and results in an efficient increase in N fixation, although there are large variations between strains of rhizobia (Mínguez et al., 1993; Sanginga et al., 1997; Tewari et al., 2004; Denton et al., 2017). Using a fixed parameter ($N_{maxfixpot}$, Eq. 9) to represent all inoculation situations such as in a global uniform calibration cannot reflect this variability. In addition, due to the difficulties in measuring both nodules and roots in the field directly, in many studies the observed BNF rates were determined from plant aboveground biomass. Excluding the root contribution to the whole plant BNF rates most likely results in an underestimation of N fixation (Córdova et al., 2019, 2020): N associated with nodules and roots in soybean and faba bean may account for 20 %–40 % of the total N accumulation at the mid-flowering phase (Unkovich and Pate, 2000; Khan et al., 2003).

Compared to non-BNF (i.e., non-nodulation treatment, see Sect. 3.1.3), BNF in LPJ-GUESS greatly improves simulated soybean yield and aboveground N mass, with an overall increase in both variables of 30 %–50 % (Table 2). Córdova et al. (2019) found a yield increase of 150 % in response to nodulation in an unfertilized treatment, but that increase was reduced to 55 % – similar to our modeled yield increase – when a high N input was applied (i.e., 135 kg N ha⁻¹). N fixation can help grain legumes to dramatically enhance their total N accumulation and to achieve higher N concentration in seeds. However, these benefits are accompanied by an increase in respiration cost amounting to 4 %–16 % of fixed total photosynthetic carbon (Kaschuk et al., 2009, 2010). Such a respiratory photosynthate consumption would reduce productivity if the photosynthesis rate was not increased to compensate for the cost. In LPJ-GUESS, as described in Sect. 2.3, we assumed that up to 50 % of daily NPP can be consumed to fix N. This approach has the advantage that legumes are able to maximize photosynthetic gain due to reduced N limitation in carboxylation capacity (V_{max}), but it entails the risk of lower productivity if too much NPP is invested in fixation. Nevertheless, in most cases our modeled NPP

cost over the soybean growing season ranged from 1 %–40 % at the site scale (Fig. S7) and 5 %–25 % on a large region (Fig. S8). Such NPP consumption was not only lower than our assumed upper limit of 50 %, but also appropriately consistent with the reported range of 14 %–32 % described by Kaschuk et al. (2009), demonstrating that the C cost scheme implemented for N fixation in our model is reasonable. Taken together, the modeled C profits due to N fixation can be attributed to the positive feedback between BNF and photosynthesis in LPJ-GUESS: C-cost-based N fixation results in a higher rate of photosynthesis because of the enhanced leaf N concentration; in turn, the increased rate compensates for the C cost, allocates more assimilate to roots, and thus enhances N fixation.

4.2 Global yields, N fixation, and %Ndfa

Agreement between FAO-reported and simulated yields at the country level was reasonable for the major soybean-producing countries. However, in some arid and semi-arid countries, the modeled yields were up to 70 % lower than FAO-reported values, probably because of the simulated low N fixation rate caused by severe water constraints (Fig. S5). By contrast, LPJ-GUESS produced an overestimation of 100 %–300 % in yield production among some African countries, with BNF rates of 300–350 kg N ha^{-1} yr^{-1} being modeled in these regions (Fig. 10a). More recent studies that report data from African farms have indicated that the soybean N fixation rate can be as low as 0–50 kg N ha^{-1} yr^{-1} in most farmers' fields, largely because of the inconsistent effectiveness of inoculation in the acid soils (Ulzen et al., 2016; Muleta et al., 2017; Vanlauwe et al., 2019). The BNF implementation and soil representation in LPJ-GUESS do not account for inoculation effectiveness in response to soil pH.

In our simulations, the annual amount of N fixed by global grain legumes (i.e., soybean and all pulses) of 17.2 ± 2.9 Tg averaged over the period 1981–2016 agreed well with the estimate of 19.4 Tg provided by Herridge et al. (2008), who used crop production statistics from FAOSTAT and legume-specific %Ndfa from farmers' fields for estimating global N fixation. In an earlier study, a total of 10 Tg N (range of 8–12 Tg N) was estimated from legume crop BNF annually (Smil, 1999), which is far lower than our findings. The discrepancy between the estimates in Smil (1999) and Herridge et al. (2008) likely reflects the lower values of %Ndfa for soybean and pulses used for calculations in Smil (1999). Also, Smil (1999) excluded belowground fixed N associated with roots and nodules, which contributes to the low estimate. Our modeled N fixation from grain legumes amounts to ∼ 12 % of the annual mean of ca. 140 Tg N that was estimated to be fixed in all global terrestrial ecosystems (Cleveland et al., 1999, 2013; Galloway et al., 2004; Wang and Houlton, 2009; Vitousek et al., 2013; Meyerholt et al., 2016; Xu-Ri and Prentice, 2017; Yu and Zhuang, 2020; Davies-Barnard

and Friedlingstein, 2020), indicating the importance of BNF input in agricultural systems for the global terrestrial N cycle, although a large proportion of the fixed N is removed in grains from the ecosystems each year.

Currently, three environmental factors, soil temperature, moisture, and soil mineral N concentration, affect modeled N fixation. As discussed in Sect. 4.1, increased soil N availability would depress N fixation as plant total N can be met more "cheaply" via soil mineral N uptake. This effect can also be seen from the spatial pattern of %Ndfa in the northern temperate region, such as the United States, western Europe, and China. Here, anthropogenic N deposition, together with the intensive application of fertilizers, results in soils being N-rich, inhibiting simulated BNF. This could explain why our modeled soybean N fixation rate was not high in East Asia and only contributed 45 ± 3 % of plant total N uptake (Table 3). In comparison, the high rate of N fixation found in tropical regions is primarily due to their high nitrogenase activity under warm and moist soil conditions (Fig. S6), resulting in %Ndfa of ∼ 70 % being modeled for all grain legumes in the tropics (i.e., Africa and Southeast Asia; Table 3). A similar spatial variation between temperate and tropical regions in N fixation was also reported by other modeling studies in global terrestrial ecosystems (e.g., Wang and Houlton, 2009; Meyerholt et al., 2016; Xu-Ri and Prentice, 2017; Yu and Zhuang, 2020). Taken together, these results reveal that LPJ-GUESS broadly captures how N management practices and climate variation affect soil N uptake and biological N fixation in grain legumes at large spatial scales.

4.3 Modeling challenges and future work

Similar to most ecosystem and crop models, specific leaf area (SLA) in LPJ-GUESS is used to compute leaf area index (LAI) and indirectly affects the amount of photosynthesis. SLA also further impacts plant total N uptake since the N demand in plant organs is always associated with the photosynthetic assimilate in the model. The disagreements between modeled and observed C–N variables in the seasonal pattern (Figs. 6–7) can therefore be partially attributed to the static value of SLA implemented in LPJ-GUESS. Some studies have shown that SLA varies with crop growth development (Boote et al., 2002; Ainsworth et al., 2007) and environmental conditions (Poorter et al., 2009). In addition, low temperature, excess radiation, water deficit, or rising CO_2 concentration would also result in reduced SLA through affecting leaf area expansion and internode elongation (Ainsworth and Long, 2005; Yin and Struik, 2010). Applying SLA as a constant in the model (see Sect. 2.4.1) cannot reflect these responses. Incorporation of dynamic SLA over the crop growing season and its response to the environment remains to be taken into account in future model development.

Despite many experimental studies on the limitation of soil water deficit in biological N fixation, the nature of the relationship between legume BNF and soil water content is

not well characterized in models. A linear water-limitation function incorporated in LPJ-GUESS (Eq. 11) implies, for instance, that the model has little potential to represent the situation when plants experience stress from excessive water (flooding). The impact of excess soil water on legume N fixation is either omitted or oversimplified in most crop models. For instance, a simple assumption adopted in Sinclair's model is that the N fixation process is stopped forcibly when flooding takes place (Sinclair et al., 1987). In STICS, the N fixation inhibition by water excess is represented as stress from hypoxia in the roots (Brisson et al., 2003). The process of legume BNF restraint by flooding is implemented into CROPGRO (Boote et al., 2008) by calculating the proportion of water-filled pore space. N fixation is assumed to only be restricted when all pore space is filled with water; however, this rule has not been well evaluated so far.

Although high soil mineral N concentration suppresses legume root nodulation and further impacts N fixation (Xia et al., 2017; Mourtzinis et al., 2018; Brar and Lawley, 2020), a moderate level of soil N in the vegetative growth stage is conducive to root growth and nodule formation, stimulating N fixation (Waterer and Vessey, 1993; Salvagiotti et al., 2008). In the field trials a specific threshold of soil N concentration above (below) which N fixation is inhibited (stimulated) is hard to measure. In addition, the timing of N application remains a challenge (Córdova et al., 2020). Some studies reported that applying N fertilizer at planting as starter N can increase yield gains because of sufficient soil-available N to stimulate early season soybean growth (Pikul et al., 2001; Osborne and Riedell, 2011; Gai et al., 2017). However, other studies argued that the best time to apply additional N would be at early reproductive growth stages, during which legumes have the greatest N demand for seed development; also, soil N reserves are depleting and N fixation rate starts slowing down (Mourtzinis et al., 2018; Córdova et al., 2019; Zhou et al., 2019). Unfortunately, as mentioned earlier, there are no consistent results on these measured factors, resulting in the difficulties in incorporating the mechanistic processes or setups into LPJ-GUESS at this point.

Taken together, the challenge of modeling legume N fixation is primarily due to its large variance between species, sites, and managements. Symbiotic nitrogen fixation by rhizobia is an extremely complex natural process, which is associated not only with host plant and soil N status in the macro-environment (see Fig. 2), but also with the process of *Rhizobium* or *Bradyrhizobium* bacteria in root nodules in the micro-environment (Rice et al., 2000). It is difficult to incorporate these two different but highly related processes into one model (Liu et al., 2011; Chen et al., 2016). Furthermore, there is an inadequate amount of information available to establish a reliable relationship between BNF and other factors such as soil pH (Rice et al., 2000; Vanlauwe et al., 2019), inoculation effectiveness (Tewari et al., 2004; Denton et al., 2017; Liu et al., 2019), salinity (Zahran, 1999; Bruning and Rozema, 2013), oxygen (Jiang et al., 2021), and other nutrition availability (Le Roux et al., 2009; Singh et al., 2012), which are currently missing from LPJ-GUESS and other crop models despite many field experiments demonstrating their importance.

5 Conclusions

In this study we implemented a mechanistic process of symbiotic biological N fixation in grain legumes into the crop module of LPJ-GUESS. The modeled C–N variables of soybean and faba bean were extensively evaluated with observed data from the site scale to a larger region. Our results showed that the BNF scheme adopted in LPJ-GUESS realistically responded to water and N managements, as well as to climate variation, and produced N fixation and yields which generally agreed with measurements.

Our model estimated that global biological N fixation in grain legumes (i.e., soybean and all pulses) was $17.2 \pm 2.9 \, \mathrm{Tg \, N \, yr^{-1}}$ during the period 1981–2016 and that the highest fixation rate occurred in tropical and temperate regions with a warm and moist climate. Soil water and temperature were dominant controls on N fixation, in addition to N fertilizer application rate. Processes missing from the model, such as inoculation effectiveness and soil acidity, might have biased our estimates of N fixation and yields at a global scale.

The dynamic process of N fixation with a C–N allocation scheme for crops in LPJ-GUESS provides an opportunity to estimate the changes in global grain legume production and global terrestrial C and N pools under future land-use or climate change scenarios. It can also help to predict and detect the potential contribution of N-fixing plants as "green manure" to reducing or removing the use of N fertilizer in global agricultural systems, considering different climate conditions, management practices, and land-use change scenarios.

References

Abu-shakra, S. S., Phillips, D. A., and Huffaker, R. C.: Nitrogen Fixation and Delayed Leaf Senescence in Soybeans, Science, 199, 973–975, 1978.

Ainsworth, E. A. and Long, S. P.: What have we learned from 15 years of free-air CO2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2, New Phytol., 165, 351–371, https://doi.org/10.1111/j.1469-8137.2004.01224.x, 2005.

Ainsworth, E. A., Rogers, A., Leakey, A. D. B., Heady, L. E., Gibon, Y., Stitt, M., and Schurr, U.: Does elevated atmospheric $[CO_2]$ alter diurnal C uptake and the balance of C and N metabolites in growing and fully expanded soybean leaves?, J. Exp. Bot., 58, 579–591, https://doi.org/10.1093/jxb/erl233, 2007.

Becker, M., Ladha, J. K., and Ali, M.: Green manure technology: Potential, usage, and limitations. A case study for lowland rice, Plant Soil, 174, 181–194, 1995.

Boote, K. J., Mínguez, M. I., and Sau, F.: Adapting the CROPGRO legume model to simulate growth of faba bean, Agron. J., 94, 743–756, https://doi.org/10.2134/agronj2002.7430, 2002.

Boote, K. J., Hoogenboom, G., Jones, J. W., and Ingram, K. T.: Modeling Nitrogen Fixation and Its Relationship to Nitrogen Uptake in the CROPGRO Model, in: Book Quantifying and Understanding Plant Nitrogen Uptake for Systems Modeling, CRC Press, 34 pp., ISBN 9780429140532, 2008.

Bouniols, A., Cabelguenne, M., Jones, C. A., Chalamet, A., Charpenteau, J. L., and Marty, J. R.: Simulation of soybean nitrogen nutrition for a silty clay soil in southern France, Field Crop. Res., 26, 19–34, https://doi.org/10.1016/0378-4290(91)90054-Y, 1991.

Brar, N. and Lawley, Y.: Short-season soybean yield and protein unresponsive to starter nitrogen fertilizer, Agron. J., 112, 5012–5023, https://doi.org/10.1002/agj2.20378, 2020.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussiere, F., Cabidoche, Y. M., Cellier, P., Debaeke, P., Gaudillere, J. P., Henault, C., Maraux, F., Seguin, B., and Sinoquet, H.: An overview of the crop model STICS, Eur. J. Agron., 18, 309–332, 2003.

Bruning, B. and Rozema, J.: Symbiotic nitrogen fixation in legumes: Perspectives for saline agriculture, Environ. Exp. Bot., 92, 134–143, https://doi.org/10.1016/j.envexpbot.2012.09.001, 2013.

Cabelguenne, M., Debaeke, P., and Bouniols, A.: EPICphase, a version of the EPIC model simulating the effects of water and nitrogen stress on biomass and yield, taking account of developmental stages: Validation on maize, sunflower, sorghum, soybean and winter wheat, Agr. Syst., 60, 175–196, https://doi.org/10.1016/S0308-521X(99)00027-X, 1999.

Chen, C., Lawes, R., Fletcher, A., Oliver, Y., Robertson, M., Bell, M., and Wang, E.: How well can APSIM simulate nitrogen uptake and nitrogen fixation of legume crops?, Field Crop. Res., 187, 35–48, https://doi.org/10.1016/j.fcr.2015.12.007, 2016.

Ciampitti, I. A. and Salvagiotti, F.: New insights into soybean biological nitrogen fixation, Agron. J., 110, 1185–1196, https://doi.org/10.2134/agronj2017.06.0348, 2018.

Ciampitti, I. A., de Borja Reis, A. F., Córdova, S. C., Castellano, M. J., Archontoulis, S. V., Correndo, A. A., Antunes De Almeida, L. F., and Moro Rosso, L. H.: Revisiting Biological Nitrogen Fixation Dynamics in Soybeans, Front. Plant Sci., 12, 727021, https://doi.org/10.3389/fpls.2021.727021, 2021.

Cleveland, C. C., Townsend, A. R., Schimel, D. S., Fisher, H., Howarth, R. W., Hedin, L. O., Perakis, S. S., Latty, E. F., Von Fischer, J. C., Hlseroad, A., and Wasson, M. F.: Global patterns of terrestrial biological nitrogen (N_2) fixation in natural ecosystems, Global Biogeochem. Cy., 13, 623–645, https://doi.org/10.1029/1999GB900014, 1999.

Cleveland, C. C., Houlton, B. Z., Smith, W. K., Marklein, A. R., Reed, S. C., Parton, W., Del Grosso, S. J., and Running, S. W.: Patterns of new versus recycled primary production in the terrestrial biosphere, P. Natl. Acad. Sci. USA, 110, 12733–12737, https://doi.org/10.1073/pnas.1302768110, 2013.

Córdova, S. C., Castellano, M. J., Dietzel, R., Licht, M. A., Togliatti, K., Martinez-Feria, R., and Archontoulis, S. V.: Soybean nitrogen fixation dynamics in Iowa, USA, Field Crop. Res., 236, 165–176, https://doi.org/10.1016/j.fcr.2019.03.018, 2019.

Córdova, S. C., Archontoulis, S. V., and Licht, M. A.: Soybean profitability and yield component response to nitrogen fertilizer in Iowa, Agrosystems, Geosci. Environ., 3, 1–16, https://doi.org/10.1002/agg2.20092, 2020.

Corre-Hellou, G., Faure, M., Launay, M., Brisson, N., and Crozat, Y.: Adaptation of the STICS intercrop model to simulate crop growth and N accumulation in pea-barley intercrops, Field Crop. Res., 113, 72–81, https://doi.org/10.1016/j.fcr.2009.04.007, 2009.

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H., and Buontempo, C.: WFDE5: bias-adjusted ERA5 reanalysis data for impact studies, Earth Syst. Sci. Data, 12, 2097–2120, https://doi.org/10.5194/essd-12-2097-2020, 2020.

Davies-Barnard, T. and Friedlingstein, P.: The Global Distribution of Biological Nitrogen Fixation in Terrestrial Natural Ecosystems, Global Biogeochem. Cy., 34, e2019GB006387, https://doi.org/10.1029/2019GB006387, 2020.

Denton, M. D., Phillips, L. A., Peoples, M. B., Pearce, D. J., Swan, A. D., Mele, P. M., and Brockwell, J.: Legume inoculant application methods: effects on nodulation patterns, nitrogen fixation, crop growth and yield in narrow-leaf lupin and faba bean, Plant Soil, 419, 25–39, https://doi.org/10.1007/s11104-017-3317-7, 2017.

DeVries, J. D., Bennett, J. M., Boote, K. J., Albrecht, S. L., and Maliro, C. E.: Nitrogen accumulation and partitioning by three grain legumes in response to soil water deficits, Field Crop. Res., 22, 33–44, https://doi.org/10.1016/0378-4290(89)90087-7, 1989a.

DeVries, J. D., Bennett, J. M., Albrecht, S. L., and Boote, K. J.: Water Relations , Nitrogenase Activity and Root Development of Three Grain Legumes in Response to Soil Water Deficits, Field Crop. Res., 21, 215–226, 1989b.

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., and Hanasaki, N.: GSWP-2: Multimodel analysis and implications for our perception of the land surface, B. Am. Meteorol. Soc., 87, 1381–1397, https://doi.org/10.1175/BAMS-87-10-1381, 2006.

Drewniak, B., Song, J., Prell, J., Kotamarthi, V. R., and Jacob, R.: Modeling agriculture in the Community Land Model, Geosci. Model Dev., 6, 495–515, https://doi.org/10.5194/gmd-6-495-2013, 2013.

Eckersten, H., Geijersstam, L., and Torssell, B.: Modelling nitrogen fixation of pea (*Pisum sativum* L.), Acta Agr. Scand. B-S. P., 56, 129–137, https://doi.org/10.1080/09064710510003050, 2006.

Elliott, J., Müller, C., Deryng, D., Chryssanthacopoulos, J., Boote, K. J., Büchner, M., Foster, I., Glotter, M., Heinke, J., Iizumi, T., Izaurralde, R. C., Mueller, N. D., Ray, D. K., Rosenzweig, C., Ruane, A. C., and Sheffield, J.: The Global Gridded Crop Model Intercomparison: data and modeling protocols for Phase 1 (v1.0), Geosci. Model Dev., 8, 261–277, https://doi.org/10.5194/gmd-8-261-2015, 2015.

Fageria, N. K.: Green manuring in crop production, J. Plant Nutr., 30, 691–719, https://doi.org/10.1080/01904160701289529, 2007.

FAO: More people, more food, worse water? a global review of water pollution from agriculture, edited by: Mateo-Sagasta, J., Zadeh, S. M., and Turral, H., FAO and IWMI, Rome and Colombo, 10 pp., ISBN 9789251307298, 2018.

FAOSTAT: Production/Crops and livestock products, FAOSTAT [data set], available at: http://www.fao.org/faostat/en/#data/QC, last access: 19 August 2021.

Fehlenberg, V., Baumann, M., Gasparri, N. I., Piquer-Rodriguez, M., Gavier-Pizarro, G., and Kuemmerle, T.: The role of soybean production as an underlying driver of deforestation in the South American Chaco, Global Environ. Chang., 45, 24–34, https://doi.org/10.1016/j.gloenvcha.2017.05.001, 2017.

Fisher, J. B., Sitch, S., Malhi, Y., Fisher, R. A., Huntingford, C., and Tan, S.-Y.: Carbon cost of plant nitrogen acquisition: A mechanistic, globally applicable model of plant nitrogen uptake, retranslocation, and fixation, Global Biogeochem. Cy., 24, GB1014, https://doi.org/10.1029/2009gb003621, 2010.

Gai, Z., Zhang, J., and Li, C.: Effects of starter nitrogen fertilizer on soybean root activity, leaf photosynthesis and grain yield, PLoS ONE, 12, e0174841, https://doi.org/10.1371/journal.pone.0174841, 2017.

Galloway, J. N., Dentener, F. J., Capone, D. G., Boyer, E. W., Howarth, R. W., Seitzinger, S. P., Asner, G. P., Cleveland, C. C., Green, P. A., Holland, E. A., Karl, D. M., Michaels, A. F., Porter, J. H., Townsend, A. R., and Vörösmarty, C. J.: Nitrogen cycles: Past, present, and future, Biogeochemistry, 70, 153–226, https://doi.org/10.1007/s10533-004-0370-0, 2004.

Gan, Y., Stulen, I., Van Keulen, H., and Kuiper, P. J. C.: Physiological changes in soybean (*Glycine max*) Wuyin9 in response to N and P nutrition, Ann. Appl. Biol., 140, 319–329, https://doi.org/10.1111/j.1744-7348.2002.tb00188.x, 2002.

Gan, Y., Stulen, I., Van Keulen, H., and Kuiper, P. J. C.: Effect of N fertilizer top-dressing at various reproductive stages on growth, N₂ fixation and yield of three soybean (*Glycine max* (L.) Merr.) genotypes, Field Crop. Res., 80, 147–155, https://doi.org/10.1016/S0378-4290(02)00171-5, 2003.

Gelfand, I. and Robertson, G. P.: A reassessment of the contribution of soybean biological nitrogen fixation to reactive N in the environment, Biogeochemistry, 123, 175–184, https://doi.org/10.1007/s10533-014-0061-4, 2015.

Hajduk, E., Właśniewski, S., and Szpunar-Krok, E.: Influence of legume crops on content of organic carbon in sandy soil, Soil Science Annual, 66, 52–56, https://doi.org/10.1515/ssa-2015-0019, 2015.

Harris, D., Pacovsky, R. S., and Paul, E. A.: Carbon Economy of Soybean–Rhizobium–Glomus Associations, New Phytol., 101, 427–440, https://doi.org/10.1111/j.1469-8137.1985.tb02849.x, 1985.

Heilmayr, R., Rausch, L. L., Munger, J., and Gibbs, H. K.: Brazil's Amazon Soy Moratorium reduced deforestation, Nature Food, 1, 801–810, https://doi.org/10.1038/s43016-020-00194-5, 2020.

Herridge, D. F., Roughley, R. J., and Brockwell, J.: Effect of rhizobia and soil nitrate on the establishment and functioning of the soybean symbiosis in the field, Aust. J. Agr. Res., 35, 149–161, https://doi.org/10.1071/AR9840149, 1984.

Herridge, D. F., Peoples, M. B., and Boddey, R. M.: Global inputs of biological nitrogen fixation in agricultural systems, Plant Soil, 311, 1–18, https://doi.org/10.1007/s11104-008-9668-3, 2008.

Houlton, B. Z., Wang, Y. P., Vitousek, P. M., and Field, C. B.: A unifying framework for dinitrogen fixation in the terrestrial biosphere, Nature, 454, 327–330, https://doi.org/10.1038/nature07028, 2008.

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., van Vuuren, D. P., and Zhang, X.: Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6, Geosci. Model Dev., 13, 5425–5464, https://doi.org/10.5194/gmd-13-5425-2020, 2020.

Irmak, S., Odhiambo, L. O., Specht, J. E., and Djaman, K.: Hourly and daily single and basal evapotranspiration crop coefficients as a function of growing degree days, days after emergence, leaf area index, fractional green canopy cover, and plant phenology for soybean, T. ASABE, 56, 1785–1803, https://doi.org/10.13031/trans.56.10219, 2013.

Jensen, E. S., Peoples, M. B., Boddey, R. M., Gresshoff, P. M., Henrik, H. N., Alves, B. J. R., and Morrison, M. J.: Legumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries. A review, Agron. Sustain. Dev., 32, 329–364, 2012.

Jiang, S., Jardinaud, M., Gao, J., Pecrix, Y., Wen, J., Mysore, K., Xu, P., Sanchez-canizares, C., Ruan, Y., Li, Q., Zhu, M., Li, F., Wang, E., Poole, P. S., Gamas, P., and Murray, J. D.: NIN-like protein transcription factors regulate leghemoglobin genes in legume nodules, Science, 374, 625–628, 2021.

Kaschuk, G., Kuyper, T. W., Leffelaar, P. A., Hungria, M., and Giller, K. E.: Are the rates of photosynthesis stimulated by the carbon sink strength of rhizobial and arbuscular mycorrhizal symbioses?, Soil Biol. Biochem., 41, 1233–1244, https://doi.org/10.1016/j.soilbio.2009.03.005, 2009.

Kaschuk, G., Hungria, M., Leffelaar, P. A., Giller, K. E., and Kuyper, T. W.: Differences in photosynthetic behaviour and leaf senescence of soybean (Glycine max [L.] Merrill) dependent on N2 fixation or nitrate supply, Plant Biol., 12, 60–69, https://doi.org/10.1111/j.1438-8677.2009.00211.x, 2010.

Kattge, J., Bönisch, G., Díaz, S., et al.: TRY plant trait database – enhanced coverage and open access, Glob. Chang. Biol., 26, 119–188, https://doi.org/10.1111/gcb.14904, 2020.

Khan, D. F., Peoples, M. B., Schwenke, G. D., Felton, W. L., Chen, D., and Herridge, D. F.: Effects of below-ground nitrogen on N balances of field-grown fababean, chickpea, and barley, Aust. J. Agr. Res., 54, 333–340, https://doi.org/10.1071/AR02105, 2003.

Kull, O.: Acclimation of photosynthesis in canopies: Models and limitations, Oecologia, 133, 267–279, https://doi.org/10.1007/s00442-002-1042-1, 2002.

Lange, S.: Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0), Geosci. Model Dev., 12, 3055–3070, https://doi.org/10.5194/gmd-12-3055-2019, 2019.

Lange, S., Menz, C., Gleixner, S., Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Müller Schmied, H., Hersbach, H., Buontempo, C., and Cagnazzo, C.: WFDE5 over land merged with ERA5 over the ocean, Version 2.0, ISIMIP Repository [data set], https://doi.org/10.48364/ISIMIP.342217, 2021.

Le Roux, M. R., Khan, S., and Valentine, A. J.: Nitrogen and carbon costs of soybean and lupin root systems during phosphate starvation, Symbiosis, 48, 102–109, https://doi.org/10.1007/BF03179989, 2009.

Lindeskog, M., Arneth, A., Bondeau, A., Waha, K., Seaquist, J., Olin, S., and Smith, B.: Implications of accounting for land use in simulations of ecosystem carbon cycling in Africa, Earth Syst. Dynam., 4, 385–407, https://doi.org/10.5194/esd-4-385-2013, 2013.

Liu, L., Knight, J. D., Lemke, R. L., and Farrell, R. E.: A side-by-side comparison of biological nitrogen fixation and yield of four legume crops, Plant Soil, 442, 169–182, https://doi.org/10.1007/s11104-019-04167-x, 2019.

Liu, Y., Wu, L., Baddeley, J. A., and Watson, C. A.: Models of biological nitrogen fixation of legumes. A review, Agron. Sustain. Dev., 31, 155–172, https://doi.org/10.1051/agro/2010008, 2011.

LPJ-GUESS: LPJ-GUESS Ecosystem Model, Lund University, available at: https://web.nateko.lu.se/lpj-guess/, last access: 14 July 2021.

Lu, C. and Tian, H.: Global nitrogen and phosphorus fertilizer use for agriculture production in the past half century: shifted hot spots and nutrient imbalance, Earth Syst. Sci. Data, 9, 181–192, https://doi.org/10.5194/essd-9-181-2017, 2017.

Ma, J., Olin, S., Anthoni, P., Rabin, S. S., Bayer, A. D., Nyawira, S. S., and Arneth, A.: Jianyong Ma/BNF in grain legumes in LPJ-GUESS, Zenodo [data set], https://doi.org/10.5281/zenodo.5148255, 2021.

Macduff, J. H., Jarvis, S. C., and Davidson, I. A.: Inhibition of N₂ fixation by white clover (Trifolium repens L.) at low concentrations of NO₃⁻ in flowing solution culture, Plant Soil, 180, 287–295, https://doi.org/10.1007/BF00015312, 1996.

Marino, D., Frendo, P., Ladrera, R., Zabalza, A., Puppo, A., Arrese-Igor, C., and Gonzalez, E. M.: Nitrogen Fixation Control under Drought Stress. Localized or Systemic?, Plant Physiol., 143, 1968–1974, https://doi.org/10.1104/pp.106.097139, 2007.

Meena, R. S., Das, A., Singh, G., and Lal, R.: Legumes for Soil Health and Sustainable Management, Springer, Singapore, https://doi.org/10.1007/978-981-13-0253-4, 2018.

Meyerholt, J., Zaehle, S., and Smith, M. J.: Variability of projected terrestrial biosphere responses to elevated levels of atmospheric CO₂ due to uncertainty in biological nitrogen fixation, Biogeosciences, 13, 1491–1518, https://doi.org/10.5194/bg-13-1491-2016, 2016.

Mínguez, M. I., Ruiz-Nogueira, B., and Sau, F.: Faba bean productivity and optimum canopy development under a Mediterranean climate, Field Crop. Res., 33, 435–447, https://doi.org/10.1016/0378-4290(93)90164-I, 1993.

Moss, B.: Water pollution by agriculture, Philos. T. Roy. Soc. B, 363, 659–666, https://doi.org/10.1098/rstb.2007.2176, 2008.

Mourtzinis, S., Kaur, G., Orlowski, J. M., Shapiro, C. A., Lee, C. D., Wortmann, C., Holshouser, D., Nafziger, E. D., Kandel, H., Niekamp, J., Ross, W. J., Lofton, J., Vonk, J., Roozeboom, K. L., Thelen, K. D., Lindsey, L. E., Staton, M., Naeve, S. L., Casteel, S. N., Wiebold, W. J., and Conley, S. P.: Soybean response to nitrogen application across the United States: A synthesis-analysis, Field Crop. Res., 215, 74–82, https://doi.org/10.1016/j.fcr.2017.09.035, 2018.

Muleta, D., Ryder, M. H., and Denton, M. D.: The potential for rhizobial inoculation to increase soybean grain yields on

acid soils in Ethiopia, Soil Sci. Plant Nutr., 63, 441–451, https://doi.org/10.1080/00380768.2017.1370961, 2017.

Northup, B. K. and Rao, S. C.: Effects of legume green manures on forage produced in continuous wheat systems, Agron. J., 108, 101–108, https://doi.org/10.2134/agronj15.0031, 2016.

Olin, S., Schurgers, G., Lindeskog, M., Wårlind, D., Smith, B., Bodin, P., Holmér, J., and Arneth, A.: Modelling the response of yields and tissue C : N to changes in atmospheric CO_2 and N management in the main wheat regions of western Europe, Biogeosciences, 12, 2489–2515, https://doi.org/10.5194/bg-12-2489-2015, 2015.

Osaki, M.: Carbon-nitrogen interaction model in field crop production, Plant Soil, 155–156, 203–206, https://doi.org/10.1007/BF00025019, 1993.

Osborne, S. L. and Riedell, W. E.: Impact of low rates of nitrogen applied at planting on soybean nitrogen fixation, J. Plant Nutr., 34, 436–448, https://doi.org/10.1080/01904167.2011.536883, 2011.

Patterson, T. G. and Larue, T. A.: Root Respiration Associated with Nitrogenase Activity (C_2H_2) of Soybean, and a Comparison of Estimates, Plant Physiol., 72, 701–705, https://doi.org/10.1104/pp.72.3.701, 1983.

Penning de Vries, F. W. T., Jansen, D. M., ten Berge, H. F. M., and Bakema, A.: Simulation of ecophysiological processes of growth in several annual crops, Centre for Agricultural Publishing and Documentation, Wageningen, Wageningen, 1989.

Peoples, M. B., Brockwell, J., Herridge, D. F., Rochester, I. J., Alves, B. J. R., Urquiaga, S., Boddey, R. M., Dakora, F. D., Bhattarai, S., Maskey, S. L., Sampet, C., Rerkasem, B., Khan, D. F., Hauggaard-Nielsen, H., and Jensen, E. S.: The contributions of nitrogen-fixing crop legumes to the productivity of agricultural systems, Symbiosis, 48, 1–17, https://doi.org/10.1007/BF03179980, 2009.

Pikul, J. L., Carpenter-Boggs, L., Vigil, M., Schumacher, T. E., Lindstrom, M. J., and Riedell, W. E.: Crop yield and soil condition under ridge and chisel-plow tillage in the northern Corn Belt, USA, Soil Till. Res., 60, 21–33, https://doi.org/10.1016/S0167-1987(01)00174-X, 2001.

Poorter, H., Niinemets, Ü., Poorter, L., Wright, I. J., and Villar, R.: Causes and consequences of variation in leaf mass per area (LMA): A meta-analysis, New Phytol., 182, 565–588, https://doi.org/10.1111/j.1469-8137.2009.02830.x, 2009.

Portmann, F. T., Siebert, S., and Döll, P.: MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling, Global Biogeochem. Cy., 24, GB1011, https://doi.org/10.1029/2008gb003435, 2010.

Pugh, T. A. M., Arneth, A., Olin, S., Ahlström, A., Bayer, A. D., Klein Goldewijk, K., Lindeskog, M., and Schurgers, G.: Simulated carbon emissions from land-use change are substantially enhanced by accounting for agricultural management, Environ. Res. Lett., 10 124008, https://doi.org/10.1088/1748-9326/10/12/124008, 2015.

Purcell, L. C. and Sinclair, T. R.: Nitrogenase Activity and Nodule Gas Permeability Response to Rhizospheric NH3 in Soybean, Plant Physiol., 92, 268–272, 1990.

Reay, D. S., Davidson, E. A., Smith, K. A., Smith, P., Melillo, J. M., Dentener, F., and Crutzen, P. J.: Global agriculture and nitrous oxide emissions, Nat. Clim. Change, 2, 410–416, https://doi.org/10.1038/nclimate1458, 2012.

Rice, W. A., Clayton, G. W., Olsen, P. E., and Lupwayi, N. Z.: Rhizobial inoculant formulations and soil ph influence field pea nodulation and nitrogen fixation, Can. J. Soil Sci., 80, 395–400, https://doi.org/10.4141/S99-059, 2000.

Richards, F. J.: A flexible growth function for empirical use, J. Exp. Bot., 10, 290–301, https://doi.org/10.1093/jxb/10.2.290, 1959.

Robertson, M. J., Carberry, P. S., Huth, N. I., Turpin, J. E., Probert, M. E., Poulton, P. L., Bell, M., Wright, G. C., Yeates, S. J., and Brinsmead, R. B.: Simulation of growth and development of diverse legume species in APSIM, Aust. J. Agr. Res., 53, 429–446, 2002.

Ryle, G. J. A., Powell, C. E., and Gordon, A. J.: The respiratory costs of nitrogen fixation in soyabean, cowpea, and white clover: I. Nitrogen fixation and the respiration of the nodulated root, J. Exp. Bot., 30, 135–144, https://doi.org/10.1093/jxb/30.1.135, 1979.

Salvagiotti, F., Cassman, K. G., Specht, J. E., Walters, D. T., Weiss, A., and Dobermann, A.: Nitrogen uptake, fixation and response to fertilizer N in soybeans: A review, Field Crop. Res., 108, 1–13, https://doi.org/10.1016/j.fcr.2008.03.001, 2008.

Sanginga, N., Dashiell, K., Okogun, J. A., and Thottappilly, G.: Nitrogen fixation and N contribution by promiscuous nodulating soybeans in the southern Guinea savanna of Nigeria, Plant Soil, 195, 257–266, https://doi.org/10.1023/A:1004207530131, 1997.

Santachiara, G., Borrás, L., and Rotundo, J. L.: Physiological processes leading to similar yield in contrasting soybean maturity groups, Agron. J., 109, 158–167, https://doi.org/10.2134/agronj2016.04.0198, 2017.

Santachiara, G., Salvagiotti, F., Gerde, J. A., and Rotundo, J. L.: Does biological nitrogen fixation modify soybean nitrogen dilution curves?, Field Crop. Res., 223, 171–178, https://doi.org/10.1016/j.fcr.2018.04.001, 2018.

Serraj, R., Sinclair, T. R., and Purcell, L. C.: Symbiotic N_2 fixation response to drought, J. Exp. Bot., 50, 143–155, https://doi.org/10.1093/jxb/50.331.143, 1999.

Sinclair, T. R., Muchow, R. C., Ludlow, M. M., Leach, G. J., Lawn, R. J., and Foale, M. A.: Field and model analysis of the effect of water deficits on carbon and nitrogen accumulation by soybean, cowpea and black gram, Field Crop. Res., 17, 121–140, https://doi.org/10.1016/0378-4290(87)90087-6, 1987.

Singh, M., Wanjari, R. H., Dwivedi, A., and Dalal, R.: Yield response to applied nutrients and estimates of N_2 fixation in 33-year-old soybean-wheat experiment on a vertisol, Exp. Agr., 48, 311–325, https://doi.org/10.1017/S0014479712000129, 2012.

Smil, V.: Nitrogen in crop production: An account of global flows adds, Global Biogeochem. Cy., 13, 647–662, 1999.

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., and Zaehle, S.: Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model, Biogeosciences, 11, 2027–2054, https://doi.org/10.5194/bg-11-2027-2014, 2014.

Soussana, J. F., Minchin, F. R., Macduff, J. H., Raistrick, N., Abberton, M. T., and Michaelson-Yeates, T. P. T.: A simple model of feedback regulation for nitrate uptake and N2 fixation in contrasting phenotypes of white clover, Annals of Botany, 90, 139–147, https://doi.org/10.1093/aob/mcf161, 2002.

Srivastava, A. K. and Ambasht, R. S.: Soil moisture control of Nitrogen fixation activity in dry tropical casuarina plantation forest, J. Environ. Manage., 42, 49–54, 1994.

Stagnari, F., Maggio, A., Galieni, A., and Pisante, M.: Multiple benefits of legumes for agriculture sustainability: an overview, Chemical and Biological Technologies in Agriculture, 4, 1–13, https://doi.org/10.1186/s40538-016-0085-1, 2017.

Tewari, K., Suganuma, T., Fujikake, H., Ohtake, N., Sueyoshi, K., Takahashi, Y., and Ohyama, T.: Effect of Deep Placement of N Fertilizers and Different Inoculation Methods of Bradyrhizobia on Growth, N_2 Fixation Activity and N Absorption Rate of Field-grown Soybean Plants, J. Agron. Crop Sci., 190, 46–58, https://doi.org/10.1046/j.0931-2250.2003.00073.x, 2004.

Thornley, J. H. M. and Cannell, M. G. R.: Modelling the components of plant respiration: Representation and realism, Annals of Botany, 85, 55–67, https://doi.org/10.1006/anbo.1999.0997, 2000.

Tian, H., Yang, J., Lu, C., Xu, R., Canadell, J. G., Jackson, R. B., Arneth, A., Chang, J., Chen, G., Ciais, P., Gerber, S., Ito, A., Huang, Y., Joos, F., Lienert, S., Messina, P., Olin, S., Pan, S., Peng, C., Saikawa, E., Thompson, R. L., Vuichard, N., Winiwarter, W., Zaehle, S., Zhang, B., Zhang, K., and Zhu, Q.: The global N_2O model intercomparison project, B. Am. Meteorol. Soc., 99, 1231–1251, https://doi.org/10.1175/BAMS-D-17-0212.1, 2018.

Tian, H., Xu, R., Canadell, J. G., Thompson, R. L., Winiwarter, W., Suntharalingam, P., Davidson, E. A., Ciais, P., Jackson, R. B., Janssens-Maenhout, G., Prather, M. J., Regnier, P., Pan, N., Pan, S., Peters, G. P., Shi, H., Tubiello, F. N., Zaehle, S., Zhou, F., Arneth, A., Battaglia, G., Berthet, S., Bopp, L., Bouwman, A. F., Buitenhuis, E. T., Chang, J., Chipperfield, M. P., Dangal, S. R. S., Dlugokencky, E., Elkins, J. W., Eyre, B. D., Fu, B., Hall, B., Ito, A., Joos, F., Krummel, P. B., Landolfi, A., Laruelle, G. G., Lauerwald, R., Li, W., Lienert, S., Maavara, T., MacLeod, M., Millet, D. B., Olin, S., Patra, P. K., Prinn, R. G., Raymond, P. A., Ruiz, D. J., van der Werf, G. R., Vuichard, N., Wang, J., Weiss, R. F., Wells, K. C., Wilson, C., Yang, J., and Yao, Y.: A comprehensive quantification of global nitrous oxide sources and sinks, Nature, 586, 248–256, https://doi.org/10.1038/s41586-020-2780-0, 2020.

Ulzen, J., Abaidoo, R. C., Mensah, N. E., Masso, C., and Abdel-Gadir, A. A. H.: Bradyrhizobium inoculants enhance grain yields of soybean and cowpea in Northern Ghana, Front. Plant Sci., 7, 1–9, https://doi.org/10.3389/fpls.2016.01770, 2016.

Unkovich, M. J. and Pate, J. S.: An appraisal of recent field measurements of symbiotic N2 fixation by annual legumes, Field Crop. Res., 65, 211–228, https://doi.org/10.1016/S0378-4290(99)00088-X, 2000.

Unkovich, M. J., Baldock, J., and Peoples, M. B.: Prospects and problems of simple linear models for estimating symbiotic N2 fixation by crop and pasture legumes, Plant Soil, 329, 75–89, https://doi.org/10.1007/s11104-009-0136-5, 2010.

Vanlauwe, B., Hungria, M., Kanampiu, F., and Giller, K. E.: The role of legumes in the sustainable intensification of African smallholder agriculture: Lessons learnt and challenges for the future, Agr. Ecosyst. Environ., 284, 106583, https://doi.org/10.1016/j.agee.2019.106583, 2019.

Vitousek, P. M., Menge, D. N. L., Reed, S. C., and Cleveland, C. C.: Biological nitrogen fixation: Rates, patterns and ecological controls in terrestrial ecosystems, Philos. T. Roy. Soc. B, 368, 20130119, https://doi.org/10.1098/rstb.2013.0119, 2013.

Voisin, A. S., Salon, C., Jeudy, C., and Warembourg, F. R.: Symbiotic N_2 fixation activity in relation to C economy of *Pisum sativum* L. as a function of plant phenology, J. Exp. Bot., 54, 2733–2744, https://doi.org/10.1093/jxb/erg290, 2003.

Voisin, A. S., Bourion, V., Duc, G., and Salon, C.: Using an ecophysiological analysis to dissect genetic variability and to propose an ideotype for nitrogen nutrition in pea, Annals of Botany, 100, 1525–1536, https://doi.org/10.1093/aob/mcm241, 2007.

Voisin, A. S., Guéguen, J., Huyghe, C., Jeuffroy, M. H., Magrini, M. B., Meynard, J. M., Mougel, C., Pellerin, S., and Pelzer, E.: Legumes for feed, food, biomaterials and bioenergy in Europe: A review, Agron. Sustain. Dev., 34, 361–380, https://doi.org/10.1007/s13593-013-0189-y, 2014.

Volkholz, J. and Müller, C.: ISIMIP3 soil input data (v1.0), ISIMIP Repository [data set], https://doi.org/10.48364/ISIMIP.942125, 2020.

von Bloh, W., Schaphoff, S., Müller, C., Rolinski, S., Waha, K., and Zaehle, S.: Implementing the nitrogen cycle into the dynamic global vegetation, hydrology, and crop growth model LPJmL (version 5.0), Geosci. Model Dev., 11, 2789–2812, https://doi.org/10.5194/gmd-11-2789-2018, 2018.

Waha, K., Van Bussel, L. G. J., Müller, C., and Bondeau, A.: Climate-driven simulation of global crop sowing dates, Glob. Ecol. Biogeogr., 21, 247–259, https://doi.org/10.1111/j.1466-8238.2011.00678.x, 2012.

Wang, E. and Engel, T.: Simulation of phenological development of wheat crops, Agr. Syst., 58, 1–24, https://doi.org/10.1016/S0308-521X(98)00028-6, 1998.

Wang, Y. P. and Houlton, B. Z.: Nitrogen constraints on terrestrial carbon uptake: Implications for the global carbon-climate feedback, Geophys. Res. Lett., 36, L24403, https://doi.org/10.1029/2009GL041009, 2009.

Wårlind, D., Smith, B., Hickler, T., and Arneth, A.: Nitrogen feedbacks increase future terrestrial ecosystem carbon uptake in an individual-based dynamic vegetation model, Biogeosciences, 11, 6131–6146, https://doi.org/10.5194/bg-11-6131-2014, 2014.

Waterer, J. G. and Vessey, J. K.: Effect of low static nitrate concentrations on mineral nitrogen uptake, nodulation, and nitrogen fixation in field pea, J. Plant Nutr., 16, 1775–1789, https://doi.org/10.1080/01904169309364649, 1993.

Weisz, P. R., Denison, R. F., and Sinclair, T. R.: Response to drought stress of nitrogen fixation (acetylene reduction) rates by field-grown soybeans, Plant Physiol., 78, 525–530, https://doi.org/10.1104/pp.78.3.525, 1985.

Williams, C. M., King, J. R., Ross, S. M., Olson, M. A., Hoy, C. F., and Lopetinsky, K. J.: Effects of three pulse crops on subsequent barley, canola, and wheat, Agron. J., 106, 343–350, https://doi.org/10.2134/agronj2013.0274, 2014.

Wu, L. and McGechan, M. B.: Simulation of nitrogen uptake, fixation and leaching in a grass/white clover mixture, Grass Forage Sci., 54, 30–41, https://doi.org/10.1046/j.1365-2494.1999.00145.x, 1999.

Wu, L., Misselbrook, T. H., Feng, L., and Wu, L.: Assessment of nitrogen uptake and biological nitrogen fixation responses of soybean to nitrogen fertiliser with SPACSYS, Sustainability, 12, 5921, https://doi.org/10.3390/SU12155921, 2020.

Xia, X., Ma, C., Dong, S., Xu, Y., and Gong, Z.: Effects of nitrogen concentrations on nodulation and nitrogenase activity in dual root systems of soybean plants, Soil Sci. Plant Nutr., 63, 470–482, https://doi.org/10.1080/00380768.2017.1370960, 2017.

Xu-Ri and Prentice, I. C.: Modelling the demand for new nitrogen fixation by terrestrial ecosystems, Biogeosciences, 14, 2003–2017, https://doi.org/10.5194/bg-14-2003-2017, 2017.

Yin, X. and Struik, P. C.: Modelling the crop: From system dynamics to systems biology, J. Exp. Bot., 61, 2171–2183, https://doi.org/10.1093/jxb/erp375, 2010.

Yu, M., Gao, Q., and Shaffer, M. J.: Simulating Interactive Effects of Symbiotic Nitrogen Fixation, Carbon Dioxide Elevation, and Climatic Change on Legume Growth, J. Environ. Qual., 31, 634–641, https://doi.org/10.2134/jeq2002.6340, 2002.

Yu, T. and Zhuang, Q.: Modeling biological nitrogen fixation in global natural terrestrial ecosystems, Biogeosciences, 17, 3643–3657, https://doi.org/10.5194/bg-17-3643-2020, 2020.

Zaehle, S. and Friend, A. D.: Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates, Global Biogeochem. Cy., 24, GB1005, https://doi.org/10.1029/2009GB003521, 2010.

Zahran, H. H.: Rhizobium-Legume Symbiosis and Nitrogen Fixation under Severe Conditions and in an Arid Climate, Microbiol. Mol. Biol. R., 63, 968–989, https://doi.org/10.1128/mmbr.63.4.968-989.1999, 1999.

Zapata, F., Danso, S. K. A., Hardarson, G., and Fried, M.: Time Course of Nitrogen Fixation in Field-Grown Soybean Using Nitrogen-15 Methodology, Agron. J., 79, 172–176, https://doi.org/10.2134/agronj1987.00021962007900010035x, 1987.

Zhang, B., Tian, H., Lu, C., Dangal, S. R. S., Yang, J., and Pan, S.: Global manure nitrogen production and application in cropland during 1860–2014: a 5 arcmin gridded global dataset for Earth system modeling, Earth Syst. Sci. Data, 9, 667–678, https://doi.org/10.5194/essd-9-667-2017, 2017.

Zhou, H., Yao, X., Zhao, Q., Zhang, W., Zhang, B., and Xie, F.: Rapid Effect of Nitrogen Supply for Soybean at the Beginning Flowering Stage on Biomass and Sucrose Metabolism, Scientific Reports, 9, 15530, https://doi.org/10.1038/s41598-019-52043-6, 2019.

Estimating the Global Influence of Cover Crops on Ecosystem Service Indicators in Croplands With the LPJ-GUESS Model

Jianyong Ma[1] , Peter Anthoni[1] , Stefan Olin[2] , Sam S. Rabin[1,3] , Anita D. Bayer[1,4], Longlong Xia[1], and Almut Arneth[1,5]

[1]Institute of Meteorology and Climate Research-Atmospheric Environmental Research, Karlsruhe Institute of Technology, Garmisch-Partenkirchen, Germany, [2]Department of Physical Geography and Ecosystems Science, Lund University, Lund, Sweden, [3]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA, [4]OHB-System AG, Weßling, Germany, [5]Institute of Geography and Geoecology, Karlsruhe Institute of Technology, Karlsruhe, Germany

Abstract Cover crops (CCs) can improve soil nutrient retention and crop production while providing climate change mitigation co-benefits. However, quantifying these ecosystem services across global agricultural lands remains inadequate. Here, we assess how the use of herbaceous CCs with and without biological nitrogen (N) fixation affects agricultural soil carbon stocks, N leaching, and crop yields, using the dynamic global vegetation model LPJ-GUESS. The model performance is evaluated with observations from worldwide field trials and modeled output further compared against previously published large-scale estimates. LPJ-GUESS broadly captures the enhanced soil carbon, reduced N leaching, and yield changes that are observed in the field. Globally, we found that combining N-fixing CCs with no-tillage technique could potentially increase soil carbon levels by 7% (+0.32 Pg C yr^{-1} in global croplands) while reducing N leaching loss by 41% (−7.3 Tg N yr^{-1}) compared with fallow controls after 36 years of simulation since 2015. This integrated practice is accompanied by a 2% of increase in total crop production (+37 million tonnes yr^{-1} including wheat, maize, rice, and soybean) in the last decade of the simulation. The identified effects of CCs on crop productivity vary widely among main crop types and N fertilizer applications, with small yield changes found in soybean systems and highly fertilized agricultural soils. Our results demonstrate the possibility of conservation agriculture when targeting long-term environmental sustainability without compromising crop production in global croplands.

Plain Language Summary Increasing crop productivity while maintaining a healthy environment is a major challenge for global agriculture. Cover crops (CCs), mostly grown during the fallow period and plowed into in soils, are expected to improve soil fertility and crop yields while reducing chemical fertilizer use, but their overall impacts on global croplands remain unknown. This study investigates the long-term influence of cover cropping on three ecosystem service indicators across four dominant farming systems (wheat, maize, rice, and soybean) using an ecological model. We find that adoption of CCs can enhance soil carbon stocks, which would contribute to slowing climate change, and benefit environments through reducing nitrogen pollution to water bodies. Among the modeled cover crop species, legumes show higher potential in increasing cash crop yields than non-legumes, but the effect is highly dependent on the crop rotation, chemical fertilizer rate, and management duration. Our results highlight that proper implementation of legume CCs can support food security and environmental sustainability in global agricultural ecosystems.

1. Introduction

Over recent decades, both global cropland areas and synthetic nitrogen (N) fertilizer application have greatly increased to feed the growing population (FAOSTAT, 2023) but with large detrimental side-effects on the environment. Estimates suggest that the conversion of natural vegetation (e.g., forests and grasslands) to cropland has reduced soil organic carbon (SOC) stocks by 32%–36% in temperate (Poeplau et al., 2011) and 25%–30% in tropical regions (Don et al., 2011). Management intensification has also caused soil fertility decline (Lal, 2004), air pollution (Reay et al., 2012; Tian et al., 2020), and freshwater eutrophication (Moss, 2008). Gaseous N emissions from fertilizer application have increased by 46% (to 3.8 Tg N yr^{-1}) for nitrous oxide (N_2O) (Tian et al., 2020) and by 78% (to 58 Tg N yr^{-1}) for ammonia (NH_3) (L. Liu et al., 2022) over the past four decades. The SOC loss from the expansion and management of agricultural land, combined with the N loss from the intense use of

Software: Jianyong Ma, Peter Anthoni, Stefan Olin, Sam S. Rabin
Supervision: Almut Arneth
Validation: Jianyong Ma
Visualization: Jianyong Ma
Writing – original draft: Jianyong Ma
Writing – review & editing: Jianyong Ma, Peter Anthoni, Stefan Olin, Sam S. Rabin, Anita D. Bayer, Longlong Xia, Almut Arneth

fertilizer greatly contribute to greenhouse gas (GHG) emissions and accelerate global warming, while undermining sustainable food production (Lu & Tian, 2017). It is crucial to enhance cropland SOC sequestration and to reduce N losses in order to mitigate climate change while still maintaining and/or increasing agricultural production for food security (Arneth et al., 2021; Poeplau & Don, 2015; P. Smith et al., 2020).

The imbalance between carbon (C) inputs (e.g., plant residue and manure application) and outputs (e.g., through crop harvest, decomposition of residues, leaching, and soil erosion) drives SOC storage changes in croplands. The adoption of minimum soil disturbance (e.g., no- or reduced tillage) has for many years been recommended as an important strategy in conservation agriculture (CA) systems to slow down the decomposition of soil organic matter (SOM) pools (Lal, 2004). However, it has been reported that the SOC benefits of no-till farming are statistically significant only in the topsoil (0–15 cm) and decline with soil depth (Haddaway et al., 2017). In global meta-analyses (Luo et al., 2010; Powlson et al., 2014), SOC stocks under no-till cropping systems were sometimes found to be even lower than conventional tillage in the deeper soil layers (>30 cm). Increasing C inputs to the soils are thus expected to be an alternative management practice for achieving SOC enhancement (Poeplau & Don, 2015). Cover crops (CCs) are plants that mostly grow during the fallow period and are incorporated into the soils as "green manure" before sowing the subsequent main crop. Experimental evidence has indicated that planting CCs within agricultural rotations may significantly increase SOC stocks by 13.8%–17.3% over a period of up to 54 years, compared to management in which the off-season is left fallow, with a global mean sequestration rate of 0.32–0.56 Mg ha^{-1} yr^{-1} (Jian et al., 2020; Poeplau & Don, 2015). In addition to increasing organic matter inputs, CCs also are able to take up excess N from the soil and thus reduce N leaching (Nouri et al., 2022; Thapa et al., 2018; Tonitto et al., 2006), as well as to prevent the soils from the compaction and erosion that happen when soils are bare (Kaye & Quemada, 2017). Moreover, using legume CCs in particular as "green manure" has been discussed as a promising technique to maintain and/or improve soil fertility and crop production because of their capacity to fix N from the atmosphere, with the co-benefits of reducing chemical fertilizer use (Abdalla et al., 2019; Ciampitti & Salvagiotti, 2018). However, these CC effects vary widely regionally due to soil properties, climate at a location, crop management, and cover crop types (Abdalla et al., 2019; Jian et al., 2020; Marcillo & Miguez, 2017; Quemada et al., 2013).

Process-based ecological models have the potential to quantify the impacts of agricultural practices on ecosystem carbon-nitrogen (C-N) and water cycles over large geographic regions and long time periods due to their mathematical representation of vegetation and soil interactions under varying environmental conditions and management (McDermid et al., 2017; Pongratz et al., 2018). These models have been widely used to investigate soil C-N dynamics and crop yields in response to CCs in different farming systems (e.g., APSIM, Chatterjee et al., 2020; DSSAT, Salmerón et al., 2014; DNDC, Singh & Kumar, 2022; ECOSYS, Qin et al., 2023). However, compared to site-level modeling studies, an assessment of the impacts of CCs across regions or globally is still lacking, as a result of inadequate management information (e.g., spatial pattern of cover crop types) and missing or incomplete cover cropping representation in models (Porwollik et al., 2022). For large-scale C-N cycle modeling assessments, alternative agricultural practices so far have been evaluated through stylized model setups with homogenous assumptions of management intensities (Jang et al., 2021; Lutz et al., 2020; Ma, Rabin, et al., 2022; Olin, Lindeskog, et al., 2015). For example, Olin, Lindeskog, et al. (2015) used the LPJ-GUESS dynamic vegetation model to explore the impacts of CCs on SOC sequestration rate across global agricultural ecosystems, assuming that all cropland grid cells adopted the same herbaceous cover crop without symbiotic N fixation. Similarly, to realistically reflect the spatial pattern of cover cropping, a recent modeling study performed by Porwollik et al. (2022) estimated with the LPJmL dynamic vegetation model how CA globally might affect soil C-N and yields in response to non-legume CCs across four cropping systems. Their model results showed the potential of cover cropping for climate change mitigation via enhanced soil C pools, but the authors suggested that future modeling assessment for N-fixing CC cultivation would be needed since this practice is identified as one practical strategy to address the conflict between the growing needs for crop production and the associated environmental problems of N loss (Abdalla et al., 2019). To date, no study has applied process-based models globally to investigate how no-till farming and legume CCs jointly affect agricultural ecosystem services, particularly in terms of soil C sequestration, N leaching from cropland, and crop yields.

Here, we employ the process-based vegetation model LPJ-GUESS (Ma, Olin, et al., 2022; Olin, Schurgers, et al., 2015; B. Smith et al., 2014) to explore the potential contribution of herbaceous N fixers to the sustainable development of agriculture production. The objective of this study is to assess and compare the effects of two cover crop types—leguminous and non-leguminous grasses—and tillage practices on SOC stocks, N leaching

loss, and agricultural productivity across global cropping systems. These three modeled ecosystem service indicators are extensively examined with worldwide site-level observed data and compared against global-level estimates from the existing literature. We aim to quantify the temporal and spatial pattern of CC impacts and to discuss the potential of this practice for climate change mitigation and crop production enhancement under present-day climate conditions.

2. Materials and Data

2.1. Model Description

LPJ-GUESS is a process-based global vegetation model that can be used to investigate plant and soil C-N dynamics and their interactions in response to changes in environment (e.g., climate, atmospheric CO_2 levels, and N deposition) and management (e.g., crop type, N fertilizer, and harvest) through simulating individual- and patch-level plant physiological and biogeochemical processes on a daily time step (B. Smith et al., 2014). Natural vegetation implemented in the model is characterized by 12 plant functional types (PFTs), with 10 woody and two herbaceous types included. PFTs differ in their phenology, photosynthetic pathway (C_3 or C_4), growth strategy, and bioclimatic limitations. Pastures are described as the competition between C_3 and C_4 grass PFTs, with half of aboveground biomass harvested annually to represent grazing impacts (Lindeskog et al., 2013). Four crop functional types (CFTs)—two temperate C_3 crops with spring and autumn sowing dates, a tropical C_3 crop representing rice, and a C_4 crop representing maize—are simulated to represent croplands, with crop-specific differences in morphological traits, dynamic C-N allocation patterns, heat requirements for growth, and N fertilization management (Olin, Schurgers, et al., 2015). Two new CFTs (i.e., soybean and pulses) with biological N fixation (BNF) have recently been added to account for the effects of legume-based cropping systems on global terrestrial N cycle (Ma, Olin, et al., 2022). For large-scale applications, the sowing date in each grid cell depends on a set of rules driven by crop- and climate-specific characteristics, with five seasonality types represented (see Waha et al. (2012) for details). Crops are harvested annually when the dynamic potential heat units (i.e., accumulated degree-days above a base temperature for each CFT) are fulfilled (Olin, Schurgers, et al., 2015). To account for crop post-harvest losses caused by mechanical damage or poor handling conditions, a harvest efficiency of 90% is used to adjust the modeled crop yields (Lindeskog et al., 2013). At present, within-year multi-cropping systems, which are common in tropical regions, have not been implemented in the model.

Cropland management options represented in LPJ-GUESS include irrigation, tillage, crop residue retention, N fertilizer and manure application, and cover crop grasses grown between two cropping seasons. Irrigation water is estimated as the amount of plant water deficit in the model and is added to the soil automatically when crops suffer from water stress. The effect of conventional tillage on heterotrophic respiration is simulated as a tillage factor of 1.94, which modifies the decay rate of four SOM carbon pools throughout the year and accelerates the soil decomposition on agricultural lands (Chatskikh et al., 2009; Pugh et al., 2015). In the standard LPJ-GUESS setup, 75% of aboveground crop residue is removed from the fields after harvest; the rest, combined with root biomass, is assumed to enter to the soil litter pool for decomposition. Synthetic N fertilizer is added to the soil mineral N pool for plant uptake at three crop development stages, with varying application rates for each CFT (see Table A2 in Olin, Lindeskog, et al., 2015). Manure is applied as a single input to cropland at sowing to account for the time required for manure N to be made available for crops. Manure is assumed to have a C:N value of 30 and is added to metabolic and structural SOM pools for decomposition (Olin, Lindeskog, et al., 2015). A variety of cover cropping options are used in this study and are described in detail below (see Section 2.2).

C-N dynamics of the soils in LPJ-GUESS are modeled by 11 SOM pools differing in C:N ratios and resistance to decay, following the CENTURY model (Parton et al., 1993). Decomposition of SOM pools results in release of CO_2 to the atmosphere (respiration) and C and N transfers between soil pools (B. Smith et al., 2014). C input to the receiver pool drives N mineralization or immobilization, as a result of maintaining mass balance and prescribed C:N ratios of the donor and receiver pool. Net N mineralization (i.e., mineralization minus immobilization), together with industrial N fertilizer and atmospheric N deposition, determine the size of the total soil mineral N pool, which is depleted by plant N uptake, as well as by crop ecosystem N losses through N leaching and gaseous N emission on a daily time step (Wårlind et al., 2014; Zaehle & Friend, 2010). Following Parton et al. (1993), mineral N leaching in the model is proportional to soil nitrate concentration and constrained by percolation rate and soil water content. N losses through soluble organic leaching are also added in LPJ-GUESS and determined by N decreases in soil microbial SOM nitrogen pool (due to decomposition), water percolation, and soil sand fraction (Wårlind et al., 2014).

Table 1
Global Simulation Setups Representing Different Cover Crop Managements (See Section 2.3.2)

Simulation	NoCC	CC_L	CC_{NL}	$CC_L NT$
Legume cover crop	No	N-fixing C_3 grass	No	N-fixing C_3 grass
Non-legume cover crop	No	No	Competing C_3 and C_4 grasses	No
Main-crop residue retention	25%	25%	25%	25%
Manure application	Yes	Yes	Yes	Yes
Mineral N fertilizer	Yes	Yes	Yes	Yes
Tillage	Yes	Yes	Yes	No

Note. NoCC—control treatment with bare fallows; CC_L—legume cover crops; CC_{NL}—non-legume cover crops; $CC_L NT$—combined management practice with legume cover crops and no tillage.

2.2. Representation of Cover Crops

CCs implemented in LPJ-GUESS so far have been simulated as competing temperate C_3 and tropical C_4 grasses grown annually between two consecutive growing seasons of main crops, replacing bare-soil fallow periods. Cover crop grass is sown on the fifteenth day after the harvest of the main crop, starting with a seedling that has an initial C mass of 0.01 kg C m^{-2} and C:N ratio 16 (Olin, Lindeskog, et al., 2015). Daily C and N mass in grasses are allocated to root and leaf pools based on a prescribed root:shoot partitioning ratio of 2 (Sainju et al., 2017), which is dynamically adjusted depending on plant water status. In the case of water stress, root allocation is increased (i.e., root:shoot partitioning ratio > 2) to help plants overcome the water limitation, following Penning de Vries et al. (1989). Cover crop grasses on fallow cropland in the simulations do not receive any management inputs (i.e., they grow under rain-fed and unfertilized conditions). Fifteen days before planting the next main crop their shoot and root biomass are added to the surface litter and the soil metabolic/structural SOM pools, respectively, for further decomposition. At this point, interplanting CCs with main crops (i.e., two plants growing beside each other at the same time) is not implemented in the model.

To account for legume CC impacts on agricultural ecosystems, we developed a new herbaceous PFT in LPJ-GUESS based on the existing C_3 grass type (Olin, Lindeskog, et al., 2015) but with BNF processes added. As in our previous work (Ma, Olin, et al., 2022), the amount of N fixed by the BNF C_3 grass is a function of soil temperature, soil water and N availability, plant development stage, and a potential N fixation rate that is dependent on net primary productivity (NPP; see Text S1 in Supporting Information S1). The fixed N is partially transported to leaves and subsequently supports photosynthesis, resulting in additional C benefits through reducing N limitation on leaf carboxylation capacity. Since fixing N from the atmosphere requires substantial chemical energy (Ryle et al., 1979), we assume that up to 50% of daily NPP may be consumed for N fixation in the model, following the findings from previous studies (Kaschuk et al., 2009, 2010; Ma, Olin, et al., 2022). More details are provided in Supporting Information S1 and in B. Smith et al. (2014).

2.3. Experimental Setups

Our study is divided into two parts. First, we test the model's ability to reproduce the observed responses in SOC stocks, N leaching, and crop yield to N-fixing and non-N-fixing CCs at various field trial sites around the world. Next, we perform four global simulations of cover crop cultivation and tillage systems (Table 1). Our analyses focus on impacts on SOC stocks, N leaching, and crop yield, first evaluating the model results against estimates from global-level studies and statistics, then analyzing and discussing the potential contribution of CCs to environmental sustainability and food security under three CA scenarios (see Section 2.3.2 below).

Model spin-up follows the protocol in Ma, Olin, et al. (2022). In order to build up the stabilized soil C-N levels on cropland, all simulations in this study are initialized with a 500-year spin-up using atmospheric carbon dioxide (CO_2) concentration from 1901 and repeating de-trended climate from 1901 to 1930 (see Table S1 in Supporting Information S1 for data sources). During spin-up, potential natural vegetation (PNV) is simulated for the first 470 years, and then the cropland fraction linearly increases from zero to the first historic value (1901) in the last 30 years. Monthly atmospheric N deposition simulated by CCMI (NCAR Chemistry-Climate Model Initiative) from 1901 to 2014 is used and interpolated to the same resolution of the climate forcing (0.5° × 0.5°) (Tian

et al., 2018). Model input data are summarized in Table S1 in Supporting Information S1, with the specific simulation experiment setups described in detail below.

2.3.1. Model Evaluation at Site Scale

To examine the model performance, cover crop field trials that also report observations of SOC stocks, N leaching, and crop yield were collected from the existing literature using the following criteria: (a) a control treatment with bare fallows (NoCC) had to be present as part of the field-based cropping experiments. We excluded greenhouse-based and vegetable farming studies, which cannot be represented by LPJ-GUESS at present. (b) Due to the absence of intercropping systems in the model (see Section 2.2 above), we only selected field trials in which CCs were either grown during the bare fallow period or undersown in main crops. For the latter case, CCs usually coexist with the main food crops for a short while (ca. 1–3 months before the main crop is harvested); CC growth is dormant during the winter months, but continues in spring, and CC crops are then terminated several days prior to the next planting of the main crop (Valkama et al., 2015). (c) To capture the variability of the observed data, CC treatments needed to cover at least two growing seasons, with the whole plant used as green manure or mulch returning to the fields. (d) Other managements, such as N fertilizer applied to main crops, had to be the same for both control and CC treatments. Cover crop trials that substituted synthetic fertilizer with green manure were thereby excluded.

A total of 43 studies carried out at 41 different sites were compiled for evaluation. Studies investigated the effects of two cover crop functional types, that is, legumes (CC_L) and non-legumes (CC_{NL}), on soil C sequestration (12 sites), N leaching (13 sites), and crop yields (29 sites) across four cropping systems (wheat, maize, rice, and soybean) and under various water and N management practices (Figure 1) and climatic zones (Ma et al., 2023). Details for these sites—their geographic coordinates, CC and main crop types, the treatment duration, as well as field management practices—are provided in Tables S2–S4 in Supporting Information S1.

Because weather data for most study locations was not available, a gridded climate data set at 0.5° resolution from GSWP3-W5E5 (Cucchi et al., 2020; Dirmeyer et al., 2006; Lange, 2019) was used as input, choosing the grid cell where the experimental sites are located. Likewise, there was not much information on land use during the years preceding the field trials for most sites. Therefore, to maintain SOM pools in equilibrium after model spin-up, we assumed that all sites were under grassland systems from 1901 to 1905, followed by a cropland period of 1906–1910, with this 5-year alternation between grassland and cropland repeated until the field trials began. Since cropland at most sites had already been present for several years at the beginning of the CC experiment, we simulated 5 years of cropland preceding the site trials at those locations for which no other information was reported. Over the experimental period, model runs were performed according to management information reported in the literature (Tables S2–S4 in Supporting Information S1). At the moment LPJ-GUESS does not simulate the cultivation of two crops simultaneously on the same field, whereas undersown CCs in the field experiments are generally grown together with main crops at least 1–3 months (Valkama et al., 2015). To better represent the total length of the cover crop growing season in the model simulations, we adjusted the sowing date of undersown CCs (referred to as the A1 runs in Table S1 in Supporting Information S1) to 1 day after the main crop harvest (instead of the default 15) and terminated the plants 1 day before the establishment of the next primary crop. For CCs solely grown on fallow cropland (A2, Table S1 in Supporting Information S1), their planting and harvest dates were assumed to be the same as the LPJ-GUESS standard setup (see Section 2.2), following the common field practice at most sites (Duval et al., 2016; Kaspar et al., 2012; Mazzoncini et al., 2011). In addition, site-specific soil physical properties—bulk density and fractions of sand, silt, and clay—derived from the literature (Ma et al., 2023) were used as external forcing to further calculate corresponding soil water characteristics and held constant across all CC simulations.

2.3.2. Global Agricultural Ecosystem Response to Cover Cropping

In this experiment we performed simulations with four CFTs—wheat, maize, rice, and soybean—which jointly provide more than two-third of the world's food supply (FAOSTAT, 2023). To detect how CCs affect cropland ecosystem services, two cover crop types—leguminous (CC_L) and non-leguminous (CC_{NL}) grasses were assessed. An additional combined practice, with N-fixing cover crop and no tillage ($CC_L NT$), was used to represent important aspects of CA. Model outputs of these three practices were compared to a control simulation with bare fallow (NoCC), applying the simulation setup given in Table 1.

The model experiments started with a baseline simulation of the historical period (1901–2014) under NoCC management after model spin-up, using dynamic gridded climate, land use/land cover, and N fertilizer data
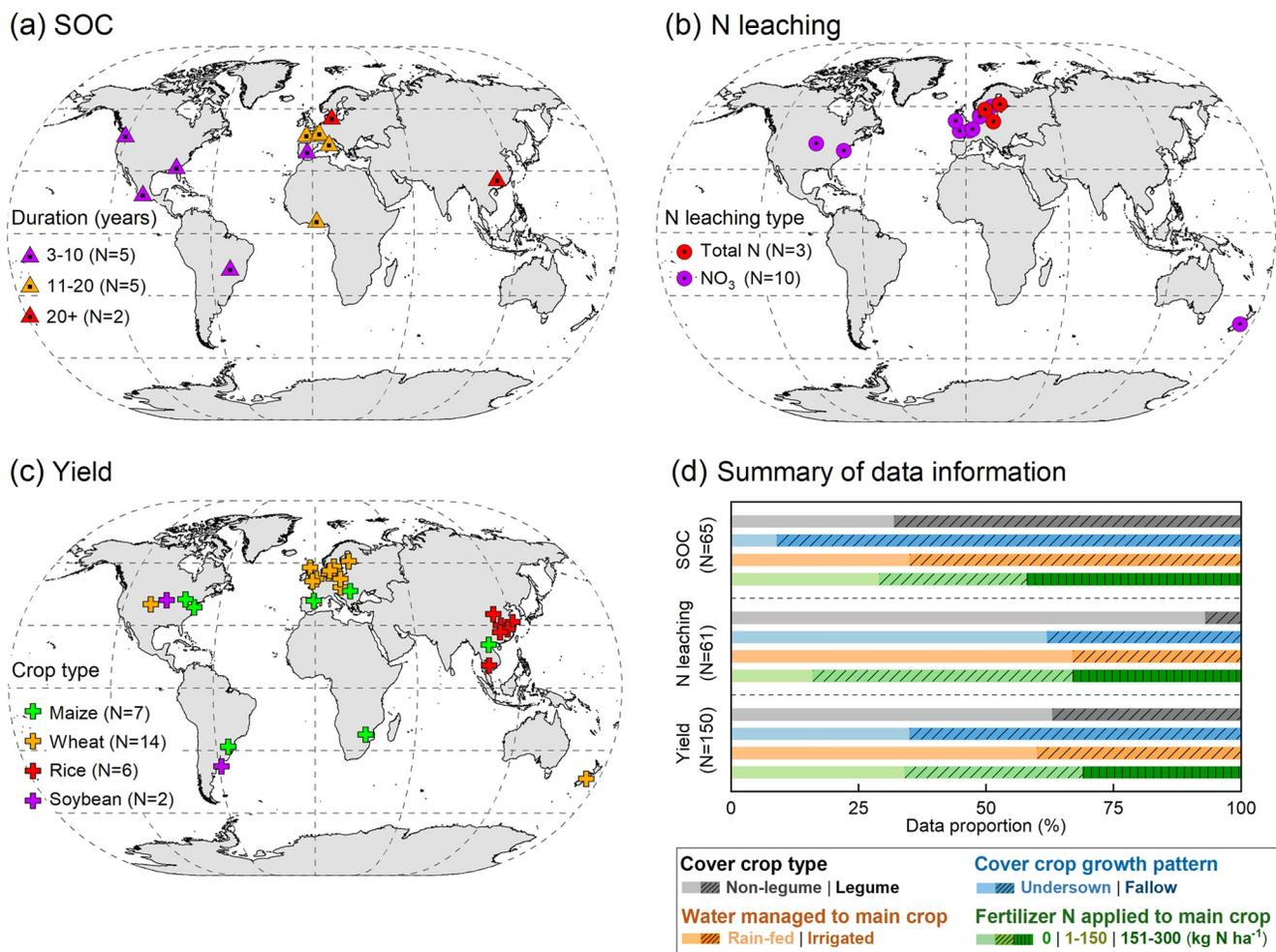
Figure 1. Distribution of cover cropping field studies used for model evaluation of cropland soil organic carbon (SOC) stocks (a), N leaching loss (b), and crop yields (c). All studied SOC sites (12) had continuously practiced cover crop (CC) cultivation for more than 3 years, and the leached N loss at the evaluated sites (13) were reported as either total N (mineral plus organic) or nitrate (NO_3). The influence of CC practice on crop production was investigated in four cropping systems (maize, wheat, rice, and soybean) at 29 sites from 16 countries. A summary of field experiments—cover crop types (legumes or non-legumes), growth patterns (undersown or fallow), and water and N fertilizer managements to main crops—is shown in (d).

($0.5° \times 0.5°$), together with atmospheric CO_2 concentration (data information described below). The result of this run was to produce present-day SOM pools on off-season fallow cropland across the globe (Table S1 in Supporting Information S1). This baseline simulation is referred to as B1.

Subsequent runs of four management practices listed in Table 1 branched from this present-day state in 2015 and are referred to as the B2 runs. These simulations ran for 36 years (the maximum duration found in cover cropping field trials in our analyzed sites; see Tables S2–S4 in Supporting Information S1) but are not intended to estimate SOC storage, N leaching and crop production through 2050; rather, they are designed to detect the relative changes in these three ecosystem indicators when replacing bare fallows with CCs. For that reason, we use constant repeated 1995–2014 climate with temperature de-trended, combined with 2014 land use, fertilizer, manure, and CO_2 concentration (Table S1 in Supporting Information S1). In order to contrast short-with long-term cover crop impacts, model outputs in the first (years 1–10) and last (years 27–36) decades were used for analysis.

For global-scale applications, LPJ-GUESS was driven by monthly mean temperature, precipitation, solar radiation, and number of wet days from the observation-based CRUJRA v2.1 data set, spanning from 1901 to 2014 at 0.5° resolution (Harris et al., 2020; Kobayashi et al., 2015). Annual atmospheric CO_2 concentration was from Meinshausen et al. (2020). Historical land use/land cover input data between 1901 and 2014 were adopted from LUH2 (Hurtt et al., 2020) and were remapped from 0.25° to 0.5° with fractions of natural vegetation, pasture, and cropland given for each grid cell. The growth distribution of various crop types, distinguishing shares of

rain-fed and irrigated crop-specific fraction per grid cell, was based on the MIRCA data set around the year 2000 (Portmann et al., 2010) and aggregated to the four CFTs simulated in this study. Thus, although the total cropland area at each grid cell varied annually over the simulation period, the relative fraction of each CFT within that cropland area remained static. To parameterize soil hydraulic properties, cropland soil texture classes in the upper soil layer (0–30 cm) from ISIMIP/GGCMI phase 3 (Volkholz & Müller, 2020) were used and held constant over the course of the model experiments. In addition, CFT-specific industrial N fertilizer and manure inputs were derived from Ag-GRID (Elliott et al., 2015) and Zhang et al. (2017), respectively, ranging from 1901 to 2014 at 0.5° resolution (Figure S1 in Supporting Information S1).

Since large-scale statistics on actual cover crop acreage do not exist, the CA area was used to represent the potential cover crop distribution on agricultural soils, following setups in a recent modeling study (Porwollik et al., 2022). We here performed all global simulations under three CA area scenarios: (a) CA_{his}, representing the approximate area of CA practice currently adopted in global croplands; (b) CA_{pot}, representing the potential agricultural lands that might implement CA systems under present socio-economic and soil biophysical conditions; (c) CA_{all}, assuming all cropland that was under CA management. Spatial pattern of CA_{his} and CA_{pot} were taken from a gridded data set developed by Porwollik et al. (2019), in which national FAO-reported CA area around the year 2005 was downscaled to grid cell level and the potential CA-suitable agricultural lands were estimated based on a range of rule-based approaches. To characterize the CA_{all} scenario, LUH2 land use data at the year 2014 were used. The spatial distribution of these three CA scenarios, as well as their total areas, are shown in Figure S2 in Supporting Information S1.

2.4. Data Analysis

Model performance at site scale was evaluated by comparing the simulated and observed ecosystem service indicators—SOC stocks, N leaching loss, and crop yield—in response to the implementation of CCs. For SOC stocks comparison, when the observed values in some field experiments were only provided as concentrations (g kg^{-1}), we converted these to stocks (Mg ha^{-1}) using Equation 1:

$$SOC_{stock} = (SOC_{con} \times BD \times D)/10 \tag{1}$$

where SOC_{stock} and SOC_{con} represent SOC stocks (Mg ha^{-1}) and concentration (g kg^{-1}), respectively. BD is bulk density (g cm^{-3}) and D is soil depth (cm).

The sampled soil depth for SOC and N leaching in our compiled data set varied from 15 to 40 cm and 60–150 cm, respectively (Tables S2–S3 in Supporting Information S1). To compare model outputs with observations, we standardized the measured SOC and N leaching from the original depth to the modeled depth of 150 cm, following the depth distribution function developed by Jobbágy and Jackson (2000) and further described by McClelland et al. (2021):

$$Y = 1 - \beta^D \tag{2}$$

$$VAR_{150} = \frac{1 - \beta^{150}}{1 - \beta^{D0}} \times VAR_{D0} \tag{3}$$

where Y is the cumulative proportion of the SOC or N leaching from the surface to depth D (cm) and β is the relative rate of decrease in these two variables with soil depth. The value of β is obtained from a meta-analysis study and set to 0.9786 for SOC and 0.9831 for N leaching (Abdalla et al., 2019). VAR denotes SOC or N leaching; D_0 is the original soil depth available in the literature; VAR_{150} and VAR_{D0} represent the cumulative SOC stocks or N leaching at 0–150 cm and original soil depth, respectively.

Based on these post-processed site-level observed data, the accuracy of the model in predicting cropland SOC stocks, N leaching, and crop yield was assessed using adjusted R^2 (the goodness of fit for the linear regression analysis), mean error (ME), mean absolute error (MAE), and the root mean square error (RMSE). In addition, to quantify the response of cropland soil C storage to CCs in comparison with the control treatment (NoCC), the annual SOC sequestration rate, ΔSOC_{rate} (Mg C ha^{-1} yr^{-1}), was calculated as:

$$\Delta SOC_{rate} = \frac{SOC_X - SOC_{NoCC}}{YR} \tag{4}$$

where SOC_X and SOC_{NoCC} are the respective SOC stocks under the cover crop and control treatments, and x denotes any cover crop practices (CC_L, CC_{NL}, and $CC_L NT$; see Table 1 for management abbreviations), and YR represents the duration (years) of management.
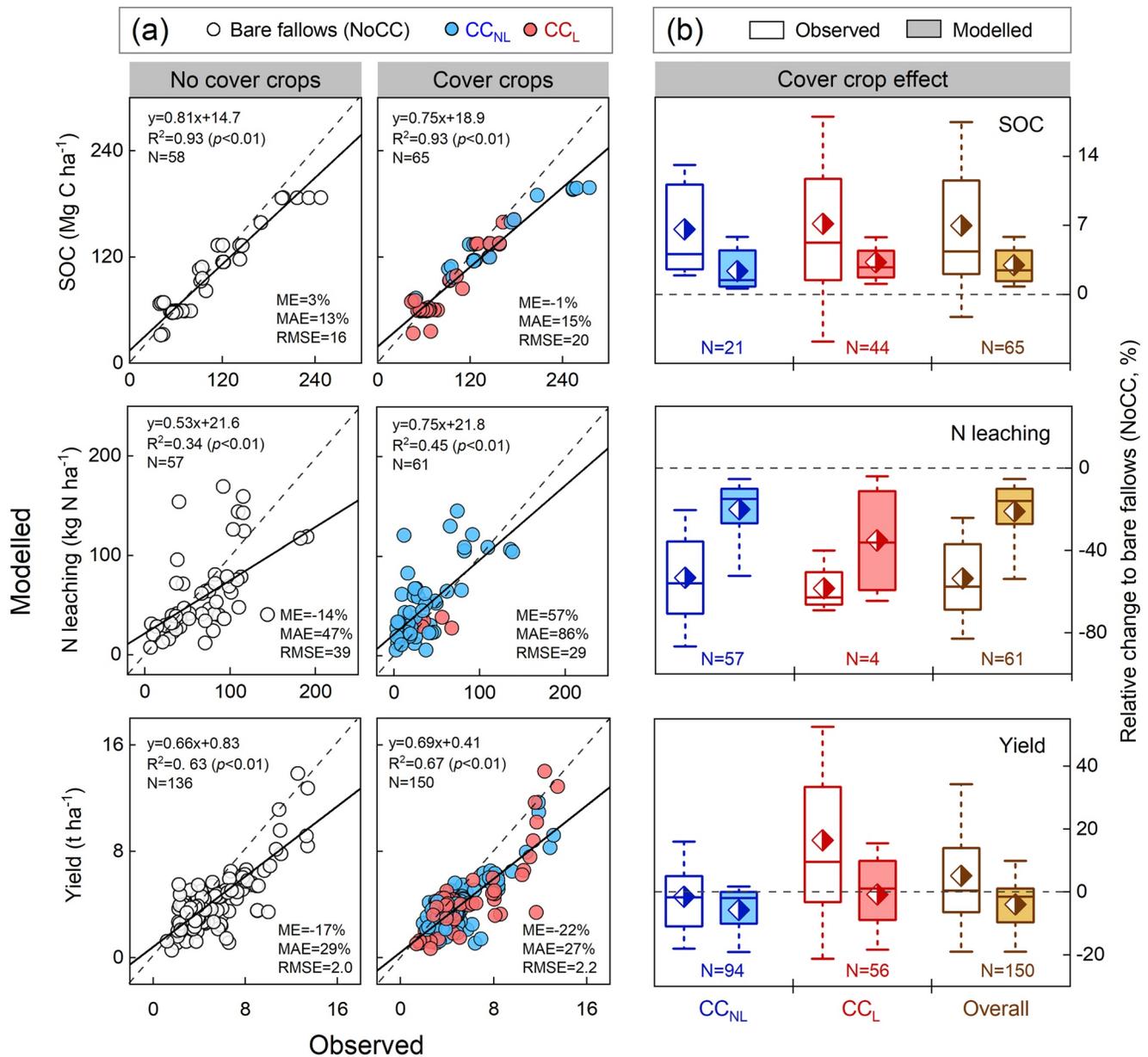
Figure 2. Comparison of modeled and observed cropland soil organic carbon (SOC) stocks, N leaching and crop yield (a) and their responses to cover crops (b) across all studied sites. The dashed line in (a) is the 1:1 line and the black bold line is a fitted linear regression; ME and MAE indicate mean error and mean absolute error, respectively (in percent); RMSE is root mean square error, with units Mg C ha⁻¹ for SOC, kg N ha⁻¹ for N leaching, and t ha⁻¹ for yield. Box plots in (b) denote the 5th and 95th percentiles by the whiskers, median and interquartile range are the box lines, and means are symbolized as diamonds. See Section 2.3.1 for treatment abbreviations and their explanations.

3. Results

3.1. Model Evaluation at Site Scale

3.1.1. Model Performance Across All Sites

Modeled SOC generally agreed well with observations, with high regression slopes (0.75–0.81) and low absolute errors (13%–15%) in the control (i.e., NoCC) and cover crop treatments (Figure 2a). We found enhanced cropland soil carbon stocks in the two simulated cover crop types compared with NoCC, indicated by positive annual SOC sequestration rates of 0.28 and 0.45 Mg C ha⁻¹ yr⁻¹ (on average) in the CC_{NL} and CC_L simulations, respectively (Table S5 in Supporting Information S1). This compared well with the observed value of 0.48 Mg C ha⁻¹ yr⁻¹

in the CC_L case, although the model underestimated the soil carbon enhancement (the range between the 5th and 95th percentiles) when all cover crop types were included (the ranges of −2.1% to 17.2% and 0.8%–5.8% for observations and simulations, respectively; Figure 2b).

Simulated N leaching from bare fallow cropland (NoCC) tended to be somewhat lower than the measurements, with a mean underestimation of 14%. By contrast, the model overestimated N losses by 57% in the cover crop experiments (Figure 2a). A positive exponential relationship between N fertilizer rate and N leaching ($p < 0.01$) was observed across a range of field sites in this study (Figure S3 in Supporting Information S1). Simulations from LPJ-GUESS mostly captured this relationship, although higher leached N rates were modeled in the highly fertilized trials (224–260 kg N ha^{-1}) compared with measurements (Figure 2a and Figure S3 in Supporting Information S1). Replacing bare fallows with CCs on average reduced N leaching by 54% in the field experiments, with the decreases ranging from 20% to 87% for non-legume types and 40%–68% for legume types (Table S5 in Supporting Information S1). LPJ-GUESS reproduced these mean differences well, but underestimated the relative changes in response to CCs, with the modeled reduction of 5%–53% and 4%–65% in the CC_{NL} and CC_L simulations, respectively (Table S5 in Supporting Information S1, Figure 2b).

In comparison with observations, LPJ-GUESS underestimated crop yields on average by 17%–22% across all field trials (Figure 2a), mainly as a result of simulated lower agricultural productivity in the unfertilized systems, particularly in wheat and rice (Figure S3 in Supporting Information S1). Compared with the bare fallows, using non-legume CCs during the off-season was modeled to reduce the subsequent main-crop production by 2%–16% across four assessed farming systems, larger than the mean observed yield reductions (1%–4%) in the field measurements. However, the implementation of N-fixing CCs in our simulations resulted in yield increases in some cases, with the production changes from −18.0% to 16.0% when all crop types were included, falling within the reported range of −21% to 52% (Table S5 in Supporting Information S1, Figure 2b). In field experiments the yield increase due to legume CCs was largest in unfertilized systems, and the impact of legume cover cropping gradually declined when main crops received high N application rates. The model reasonably reproduced the decreased trend of yield benefits to N fertilizer increases, but generally underestimated these effects in most N fertilization trials (Figure S3 in Supporting Information S1).

3.1.2. Soil Organic Carbon Response to Cover Crops

In a long-term (15 years) experiment Mazzoncini et al. (2011) tested the SOC response to agricultural management. Three cover crop treatments (NoCC, CC_{NL}, and CC_L) with two tillage strategies and four N fertilization rates were conducted in a cropping system that grew first maize, followed by wheat-maize rotation, and sunflower in the last year (Italy, 10.3°E, 43.7°N; see Table S2 in Supporting Information S1). Since the main crop—sunflower—has not been incorporated in the current version of LPJ-GUESS, we modeled this crop type as wheat aiming to test whether we could nevertheless reproduce the general response of SOC to the different managements.

After 15 years of cover cropping, the observed mean SOC stocks in the field trials increased from 92.5 to 89.7 Mg C ha^{-1} in 1993 to 97.7 and 102.3 Mg C ha^{-1} in 2008 for CC_{NL} and CC_L treatment, respectively (Mazzoncini et al., 2011). The modeled soil carbon changes, averaged across a range of management options, were 91.2–97.1 Mg C ha^{-1} in the CC_{NL} simulation and 91.2–98.6 Mg C ha^{-1} in the CC_L simulation over the same period, suggesting overall good model performance although SOC increases in the CC_L simulation were underestimated (Figure 3a). The 15-year adoption of non-legume and legume CCs was simulated to sequester 0.07 and 0.17 Mg C ha^{-1} yr^{-1} of soil carbon (ΔSOC_{rate}, Equation 4), respectively, relative to bare fallows (NoCC). The observations showed similar responses but with higher sequestration rates of 0.26 and 0.57 Mg C ha^{-1} yr^{-1} (Table S6 in Supporting Information S1). Over the experimental period there was an obvious underestimation of the simulated total aboveground biomass for all treatments. This may be partially due to the lower shoot biomass of CCs in the model experiments compared with observations (Figure 3b). Moreover, LPJ-GUESS at this point does not simulate the growth of weeds, which amount to ~10%–30% of the total aboveground dry matter in the field measurements (Figure 3b).

3.1.3. Nitrate Leaching and Crop Yield Response to Cover Crops

The ability of the model to simulate observed nitrate leaching and crop yields in response to CCs was examined using data from a 4-year field experiment carried out in a rain-fed maize-soybean rotation system in Ames, USA (93.7°W, 42.1°N; see Tables S3–S4 in Supporting Information S1). At this site, ryegrass was the overwintering cover crop (Kaspar et al., 2012), solely cultivated on fallow cropland; a legume cover cropping experiment was not conducted in the field.
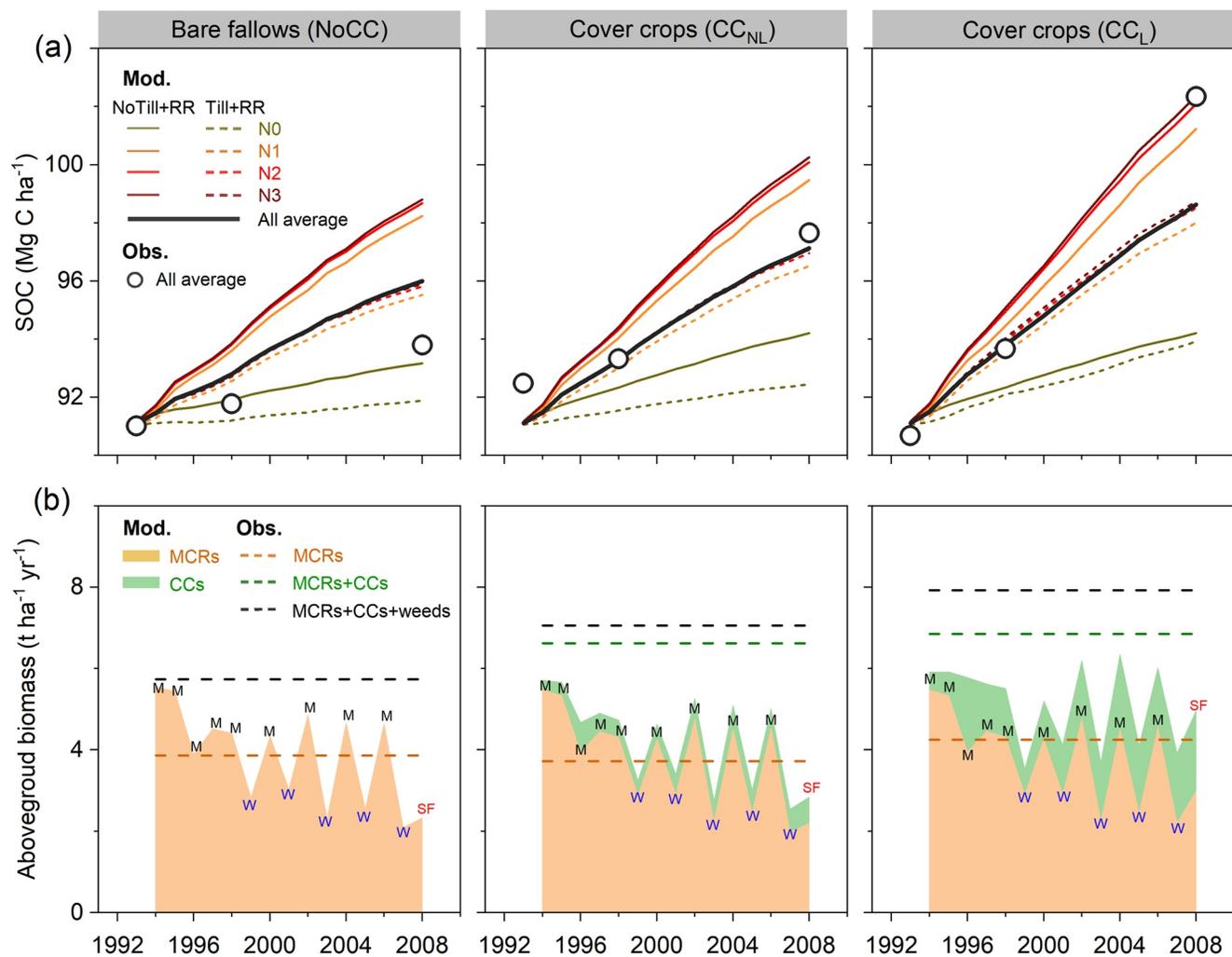
Figure 3. Modeled and observed cropland soil organic carbon (SOC) (a) and aboveground biomass (b) under three cover crop treatments at a rain-fed field site with Mediterranean climate in Pisa (Italy) between 1993 and 2008. The main crop of sunflower only planted in the last year of the field experiments was modeled as wheat. The observed SOC stocks in Mazzoncini et al. (2011) were reported as the mean values of two tillage strategies and four N fertilizer levels and were labeled as "All average" in (a). N0, N1, N2, and N3 in (a) are respectively no N, low N, medium N and high N fertilization rates, with 0, 60, 120, and 180 kg N ha^{-1} for wheat and 0, 100, 200, and 300 kg N ha^{-1} for maize. The observed aboveground biomass shown in dashed lines in (b) represents the mean values from 1993 to 2008. Abbreviations: NoTill—no tillage; Till—Tillage; RR—100% of main-crop residue retention; M—maize; W—wheat; SF—sunflower; MCRs—main crop residues; CCs—cover crops.

Shoot biomass and N mass of the C$_3$ herbaceous CCs in our simulations first increased rapidly between October and November, and then commenced again in late March in response to the increasing temperature in spring (Figure 4a). With exception of 2008, the modeled aboveground production of CCs was lower than the field measurements; differences between modeled and measured were 0.7 and 0.5 t ha^{-1} in 2007 and 2009, respectively. Over the cropping seasons there was an underestimation of maize yield for all simulations, ~15%–26% lower than the observed values of 11.2–11.7 t ha^{-1} (Figure 4a). Replacing bare fallows with CCs was simulated to reduce maize production by 6%–8%, in line with the observed loss of 1%–4%, likely reflecting indirect competition for water and nutrients between CCs and main crops. These negative impacts of CCs on yield were also found in the field-grown soybean trials (reduction of 1%–13%) but not found in our model experiments (Figure 4a).

The simulated nitrate leaching from bare fallow cropland (NoCC) ranged from 32 to 95 kg N ha^{-1} yr^{-1} during 2006–2009 (with a total cumulative loss of 219 kg N ha^{-1} until 2009; Figure 4b), and exceeding the observed values of 29–67 kg N ha^{-1} yr^{-1} over the same period (195 kg N ha^{-1} in total; Kaspar et al., 2012). Using CCs mitigated this hydrological N loss by 35%–75% to 9–37 kg N ha^{-1} in the field trials, comparable but higher than our modeled reduction of 13%–34%. The cause for the underestimated reduction in N leaching may be that the simulated soil N uptake by CCs was lower than the field measurements, given that shoot N mass of CCs was
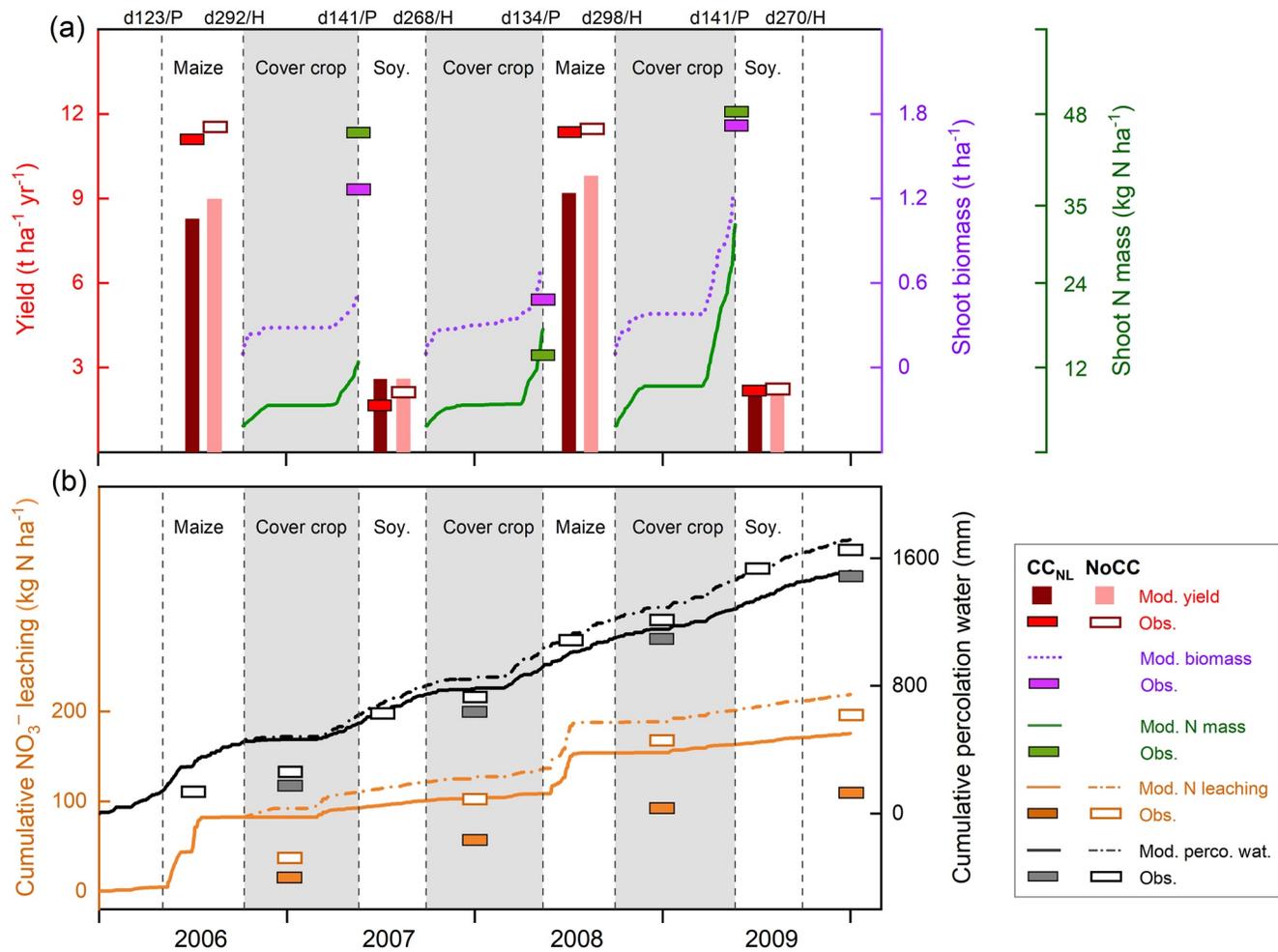
Figure 4. Modeled and observed shoot biomass and shoot N mass of cover crops, and main crop yield (a) in a rain-fed maize-soybean rotation system in Ames (USA) from 2006 to 2009, and the response of cumulative percolation water and nitrate leaching to cover crop practice compared to the control treatment with bare fallows (b). The observations from overwintering cover crops(ryegrass) reported in Kaspar et al. (2012) were chosen for model evaluation. Maize during the growing season received 225 and 198 kg N ha^{-1} of fertilizer application in 2006 and 2008, respectively, and no chemical N fertilizer was applied to soybean over the entire experimental period. Abbreviations: NoCC—control treatment with bare fallows; CC$_{NL}$—non-legume cover crops; d—day of the year; P—planting date of main crops; H—harvest date of main crops; Soy—soybean.

below the observations (Figure 4a). Step changes in simulated N leaching over the cropping seasons 2006 and 2008 (Figure 4b) corresponded to the high fertilization rates of 198–225 kg N ha^{-1} in maize systems. Such an increase was absent in 2007 and 2009 mainly because soybeans were not fertilized. In addition, the replacement of bare fallows with CCs in our simulations had the potential to reduce soil percolation water by 3%–12%, agreeing well with the observed decreases of 4%–20% (Figure 4b).

3.2. Global Crop Ecosystem Responses to Cover Crops

3.2.1. Soil Carbon Stocks

Our simulations of the three explored CC managements resulted in a net soil C increase across global croplands compared with the control treatment (NoCC), with the largest SOC sequestration rates (ΔSOC_{rate}, Equation 4) found in warm and moist regions (Figure 5 and Figure S4 in Supporting Information S1). For the 36-year simulation period, the maximum annual rates of soil C sequestration in the CC$_{NL}$ and CC$_L$ runs were reached in the sixth year after introducing cover cropping, whereas in the CC$_L$NT simulation they were already achieved in the fourth year after the implementation of altered management (Figure 5). After these initial peaks, the annual soil C accumulation effect persisted over the course of the remaining simulation period, but with declining rates. On
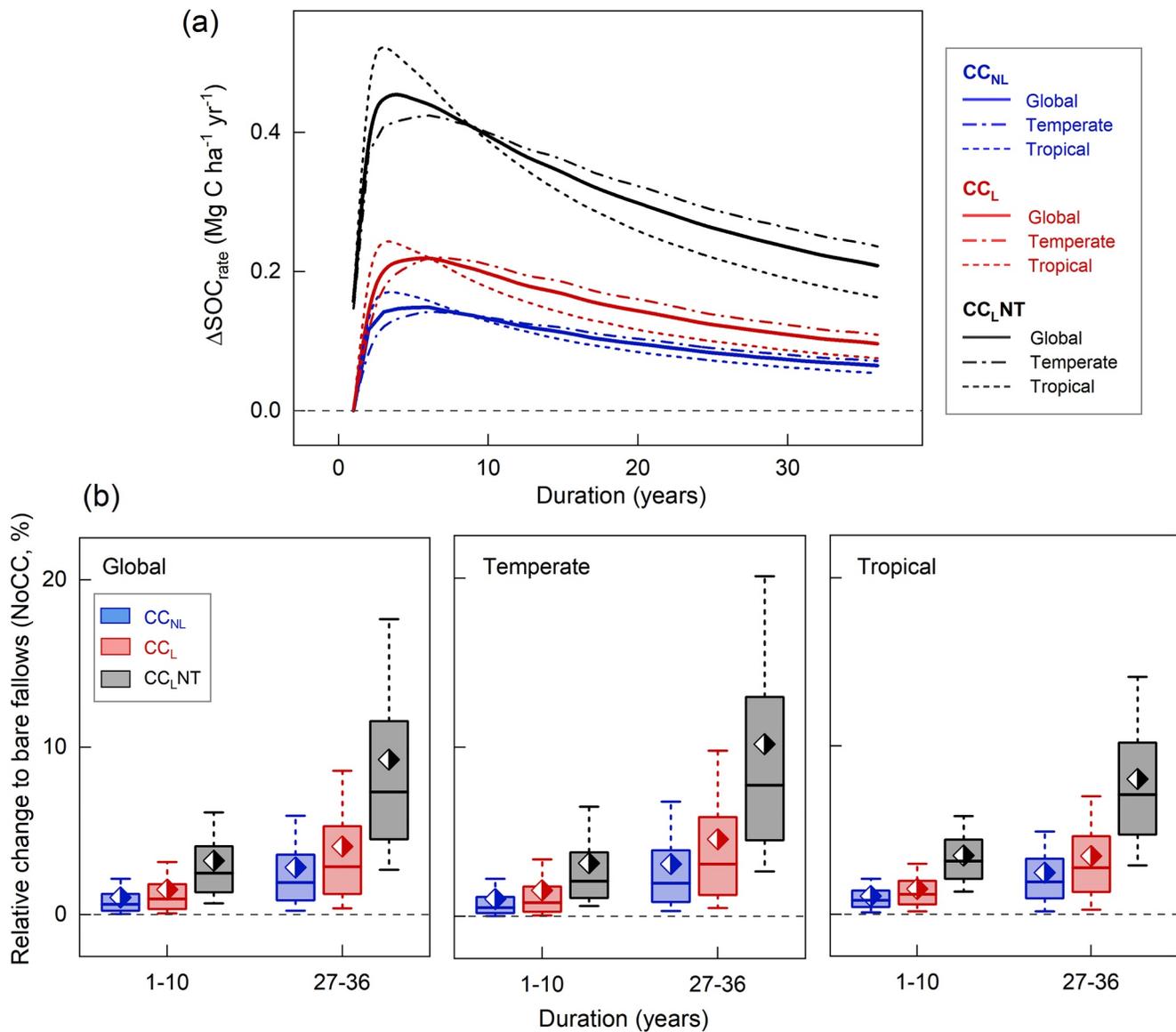
Figure 5. Area-weighted aggregated average annual soil C sequestration rate (Equation 4, Mg C ha^{-1} yr^{-1}) across global (1,597 × 10^6 ha), temperate (987 × 10^6 ha), and tropical (606 × 10^6 ha) croplands for three cover crop managements (CC$_{NL}$: blue; CC$_L$: red; CC$_L$NT: black) in the CA$_{all}$ scenario (a), and relative responses (%) of soil organic carbon stocks to these cover crop strategies compared with the control treatment (bare fallows, NoCC) in the first and last decades of the 36-year simulation period (b). The temperate region in this study is defined as the latitudes from 23.5° to 60° N/S of the equator, and latitudes between 23.5°S and 23.5°N are classified as the tropics. Box plots in (b) denote the 5th and 95th percentiles with whiskers, median and interquartile range with box lines, and mean with diamonds across all cropland grid cells (global:35,039; temperate:21,223; tropical:12,942).

average, using CCs was modeled to sequester 0.10, 0.14, and 0.32 Mg C ha^{-1} yr^{-1} of soil carbon in the CC$_{NL}$, CC$_L$, and CC$_L$NT runs, respectively (Figure 5).

Under the CA$_{all}$ scenario, modeled total soil C stocks (0–150 cm) of the various managements ranged from 164.9 to 176.4 Pg C across global croplands, somewhat larger than the published estimates for the topsoil layer (140 Pg C in 0–30 cm, Zomer et al., 2017; 115 Pg C in 0–50 cm, Ren et al., 2020) and within the reported values for the depth 0–100 cm ranging between 157 and 164 Pg C (Global Soil Data Task, 2014; Jobbágy & Jackson, 2000) and 210 Pg C (for 0–200 cm; Jobbágy & Jackson, 2000). In comparison with bare fallows (NoCC), simulations from LPJ-GUESS resulted in an increase of soil C storage by 3.8 (+2.3%) and 5.4 Pg C (+3.3%) after 36 years of implementation of non-legume (CC$_{NL}$) and legume cover crops (CC$_L$), respectively, between the main cropping seasons. Adopting no tillage (CC$_L$NT) further contributed to increasing modeled C storage by 11.5 Pg C (+7.0%) across global croplands (CA$_{all}$ scenario; Table 2).

Table 2
Modeled Total Cropland Soil Organic Carbon Stocks (0–150 cm), N Leaching Loss, and Crop Production With Alternative Cover Crop Managements Under Three CA Area Scenarios in the First and Last Simulated Decades, Compared With Literature-Based Estimates

| Scenario[a]	Management	Soil C stock, total (Pg C)		N leaching, total (Tg N yr^{-1})		Crop production[b] (million tonnes per year)	
		1–10 years	27–36 years	1–10 years	27–36 years	1–10 years	27–36 years
CA$_{his}$ (126 × 10^6 ha)	NoCC	15.8	15.6	0.88	0.80	301	287
	CC$_{NL}$	15.9	15.9	0.52	0.49	286	292
	CC$_L$	16.0	16.1	0.58	0.54	295	306
	CC$_L$NT	16.2	16.7	0.51	0.48	279	294
CA$_{pot}$ (590 × 10^6 ha)	NoCC	68.9	68.0	5.2	5.4	1,145	1,126
	CC$_{NL}$	69.4	69.5	3.3	3.2	1,068	1,119
	CC$_L$	69.6	70.2	3.7	3.6	1,105	1,172
	CC$_L$NT	70.3	72.5	3.2	3.2	1,034	1,125
CA$_{all}$ (1,597 × 10^6 ha)	NoCC	167.3	164.9	18.4	17.8	2,785	2,743
	CC$_{NL}$	168.5	168.7	10.8	10.5	2,635	2,765
	CC$_L$	169.1	170.3	12.2	11.7	2,714	2,875
	CC$_L$NT	171.1	176.4	10.7	10.5	2,557	2,780
Other studies (global cropland)		115[c]; 140[d]; 157–210[e]; 164[f]		14–20[g]; 23[h]; 26[i]; 31[j]		2806[k]	

[a]See Figure S2 in Supporting Information S1 for spatial pattern of three CA area scenarios. [b]Summed yield of four crop types: maize, wheat, rice, and soybean. [c]Ren et al. (2020), 0–50 cm, 1,667 × 10^6 ha. [d]Zomer et al. (2017), 0–30 cm, 1,631 × 10^6 ha. [e]Jobbágy and Jackson (2000), the estimate for 0–100 cm is 157 Pg C, and that for 0–200 cm is 210 Pg C, 1,400 × 10^6 ha. [f]Global Soil Data Task (2014), 0–100 cm, 1,518 × 10^6 ha. [g]Smil (1999). [h]J. Liu et al. (2010). [i]Lin et al. (2001). [j]Q. Liu et al. (2019). [k]FAOSTAT (2023); reported total production in the year 2014 were used for comparison: 1,040, 729, 731, and 306 million tonnes for maize, wheat, rice, and soybean, respectively.

3.2.2. Cropland N Leaching and Yields

In addition to soil C benefits, CCs resulted in a reduction in simulated N leaching in most global croplands (i.e., CA$_{all}$ scenario), with the largest decreases (∼75%–90%) found in Russia and large parts of Africa, regions where mineral N fertilizer application were rather low (Figure S1 in Supporting Information S1). Modeled N leaching reduction in response to CCs in China, Western Europe, and the United States—areas with substantial fertilizer application (Figure S1 in Supporting Information S1)—were still 0%–45% for the 36-year simulation period (Figure 6). Our simulated total nitrogen loss of 17.8–18.4 Tg yr^{-1} from fallow cropland (NoCC) was in good agreement with statistics-based estimates of 14–23 Tg N yr^{-1} (J. Liu et al., 2010; Smil, 1999), but lower than the findings of 26–31 Tg N yr^{-1} in Lin et al. (2001) and Q. Liu et al. (2019) who uses a modeling approach (Table 2). Replacing bare fallows with cover cropping across global croplands was modeled to reduce N leaching by 7.3–7.6 and 6.1–6.2 Tg N yr^{-1} in the CC$_{NL}$ and CC$_L$ runs, respectively. The latter (i.e., CC$_L$) was ∼17% lower than the decreases of 7.3–7.7 Tg N yr^{-1} from CC$_L$NT (Table 2, Figure 6), supporting arguments for practicing conservation tillage techniques to mitigate hydrological N losses.

The modeled impacts of legume cover crops (CC$_L$) on yields of the main crops showed large spatial variation (Figure 7; see Figure S5 in Supporting Information S1 for the spatial patterns of CC$_{NL}$ and CC$_L$NT). Small, and inconclusive with respect to their direction, yield changes between −5% and 5% (36-year average) were found in China across all crop types, likely as a consequence of the high N fertilizer input (Figure S1 in Supporting Information S1). A widespread yield loss in response to CCs was seen in northern cold and temperate dry climates, whereas yields in humid regions—such as the eastern USA, southern China, and most of South America and Africa—increased (Figure 7), reflecting high biomass and high N fixation rates (Figure S6 in Supporting Information S1). However, these modeled impacts varied widely between different cropping systems, with the largest yield variability found in maize and wheat, followed by rice. Productivity of soybean crops responded only little to legume CCs (Figure 8).

Our model simulations under bare fallow management (NoCC) resulted in a total crop production of 2,743–2,785 million tonnes yr^{-1} globally, consistent with FAO-reported estimate of 2,806 million tonnes in the year 2014 (Table 2), implying the reliability of the current model version to reproduce food production at the global scale.
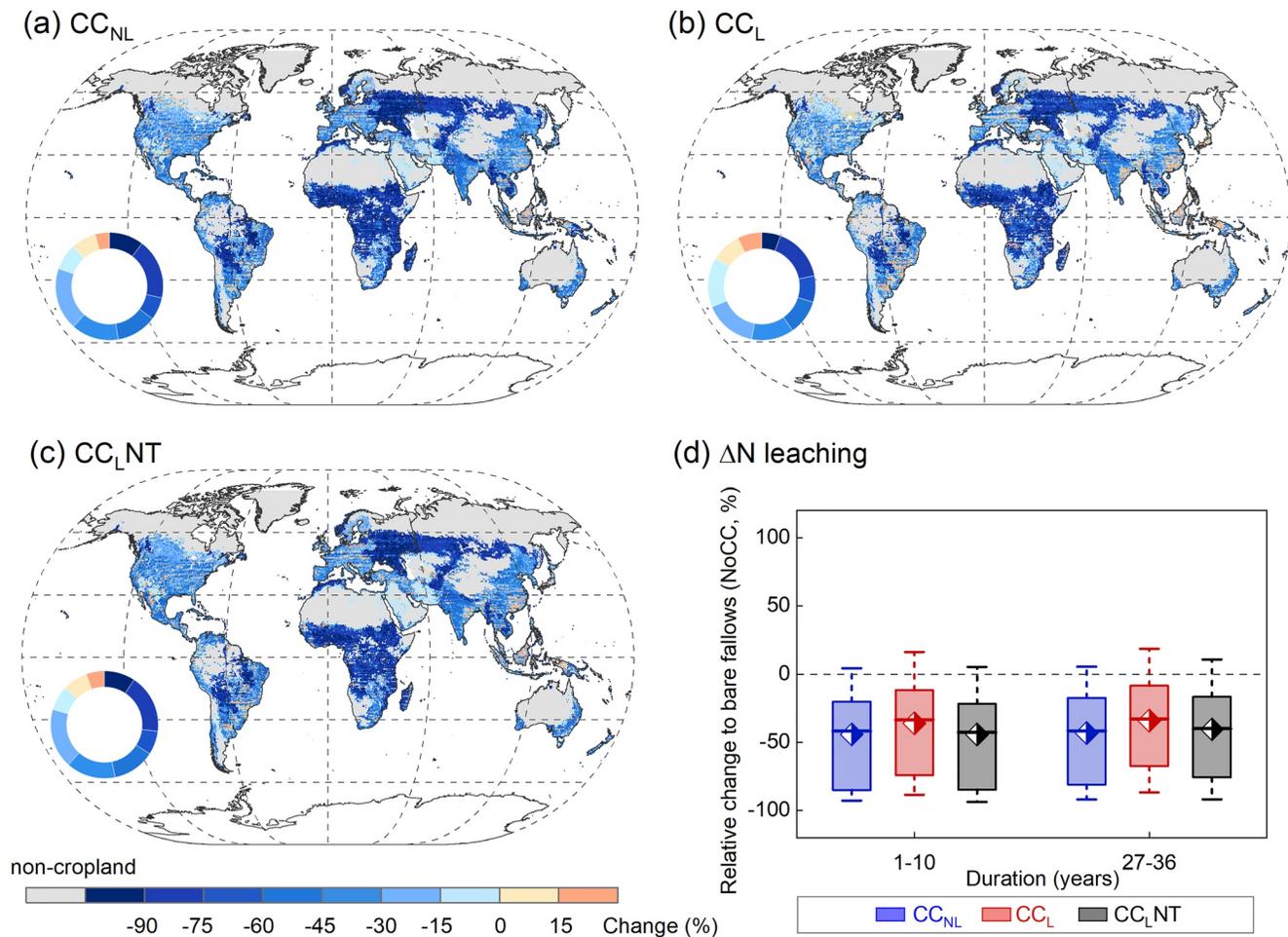
Figure 6. Maps of the simulated responses (%, 36-year average) of cropland N leaching to CC_{NL} (a), CC_L (b), and CC_LNT (c) managements, relative to the control treatment with bare fallows (NoCC) in the CA_{all} scenario. Box plots of these responses in the first and last simulated decades are shown in (d), denoting the 5th and 95th percentiles with whiskers, median and interquartile range in box lines, and mean with diamonds across all cropland grid cells (35,039). The inset donut plots in (a–c) represent the area proportion of each classified ΔN leaching from the total cropland area.

Compared with fallow soils during off-season period, using CCs was modeled to potentially reduce main-crop yield in the first decade for the 36-year simulation, with mean decreases of 6%, 3%, and 8% in CC_{NL}, CC_L, and CC_LNT, respectively. However, these negative yield effects were gradually diminished over the course of simulation, and turned to positive impacts in the last decade, with slight production increases of 1%–5% simulated for the three assessed managements in comparison with the control treatment (Table 2).

4. Discussion

4.1. Soil Carbon Stocks

LPJ-GUESS simulates cropland soil carbon stocks across all the evaluated sites well, although the measured SOC increase in response to CCs is generally underestimated (Figure 2). One likely explanation for this discrepancy is the low biomass production of CCs in the model experiments (Figures 3 and 4), resulting in less C input to the soil pools compared to the field measurements. Experimental evidence from the field sites has shown that the amount of biomass C added to the soil through CCs varies widely between cover crop species (Constantin et al., 2010; Kuo et al., 1997; Sainju et al., 2002). Using two grass functional types (i.e., groups of grasses with similar functional behaviors; see Section 2.2) to represent all cover crop situations in our standardized evaluation cannot reflect this variability. Also, when comparing herbaceous CC effects on soil carbon stocks, belowground C input via roots has been proven to stably enhance SOC sequestration in the field measurements (Blanco-Canqui
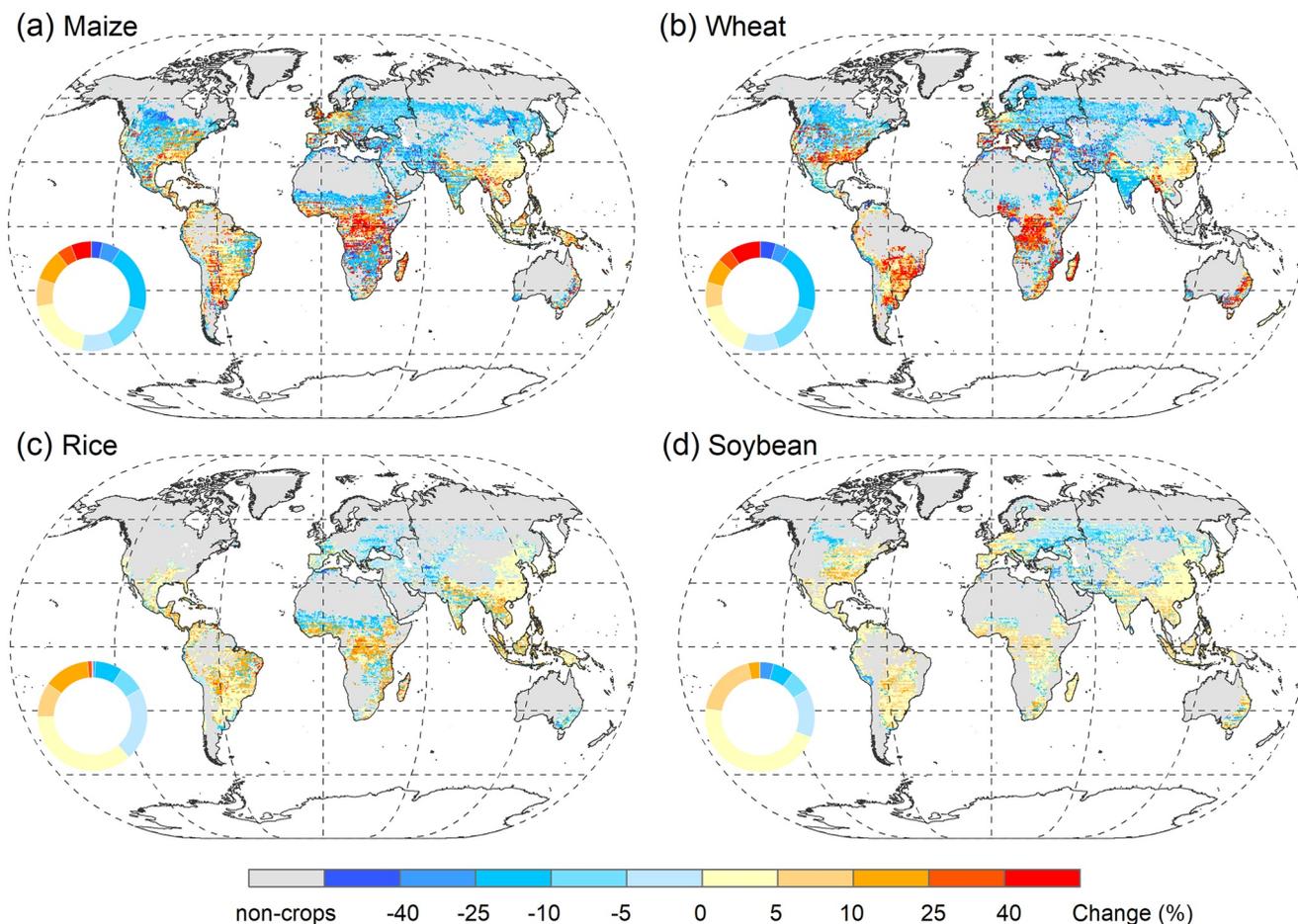
Figure 7. Maps of simulated crop-specific production in response to legume cover crops (CC_L) compared to the control treatment with bare fallows (NoCC) in the CA_{all} scenario: maize (a), wheat (b), rice (c), and soybean (d). Modeled crop-specific production at each grid cell is calculated as the area-weighted aggregated results in rain-fed and irrigated conditions. Global total cropping areas (rain-fed and irrigated) in the year 2014 used in this study are 184.5, 247.7, 151.7, and 95.9×10^6 ha for maize, wheat, rice, and soybean, respectively, with rain-fed proportions of 84%, 77%, 44%, and 94% for those four crop types. Yield relative changes (%) in maps are given as the mean values for the 36-year simulation period. The inset donut plots represent the area proportion of each classified Δyield from the total crop-specific area.

et al., 2015; Rasse et al., 2005) but with large variability due to differences in soil types, local climate, and CC species (Sainju et al., 2017). For instance, in a 2-year U.S. trial, Kuo et al. (1997) found that the root-to-shoot ratio of plant biomass C grown under natural conditions ranged from 0.5 to 0.8 for ryegrass (non-legume) and 0.2–0.5 for hairy vetch (legume). In comparison, higher root-to-shoot ratios in perennial grasses (e.g., intermediate wheatgrass and smooth bromegrass) ranging from 1.0 to 3.5 were reported in another U.S. field experiment with continental climate, depending on soil sampling depth and nutrient availability (Sainju et al., 2017). Here, we implemented a prescribed root-to-shoot ratio of 2.0 to broadly represent below- and aboveground biomass productions in herbaceous plants based on literature values (see Section 2.2). Whether this set value affects the simulated root-derived carbon from CCs is difficult to assess because root biomass information was typically unavailable from the test sites. In addition, at this point LPJ-GUESS does not account for potential C inputs through weeds (Figure 3; Mazzoncini et al., 2011), which may further bias our assessed CC effects on SOC sequestration rates at site scale.

Our modeled global-scale small SOC increase of 1.0%–2.8% for non-legume cover crops (CC_{NL}) and 1.5%–4.1% for legumes (CC_L) (Figure 5) agreed with the meta-analysis of Abdalla et al. (2019) and Poeplau and Don (2015), in which replacing bare fallows with CCs statistically showed no significant difference between cover crop types for SOC sequestration, with a mean increase of 4.1% and 4.5% found for non-legumes and legumes, respectively. However, these reported impacts were somewhat lower than a recent synthesis conducted by Jian et al. (2020), who found that cover cropping would result in a net SOC sequestration of 0.56 Mg C ha^{-1} yr^{-1} with all cover
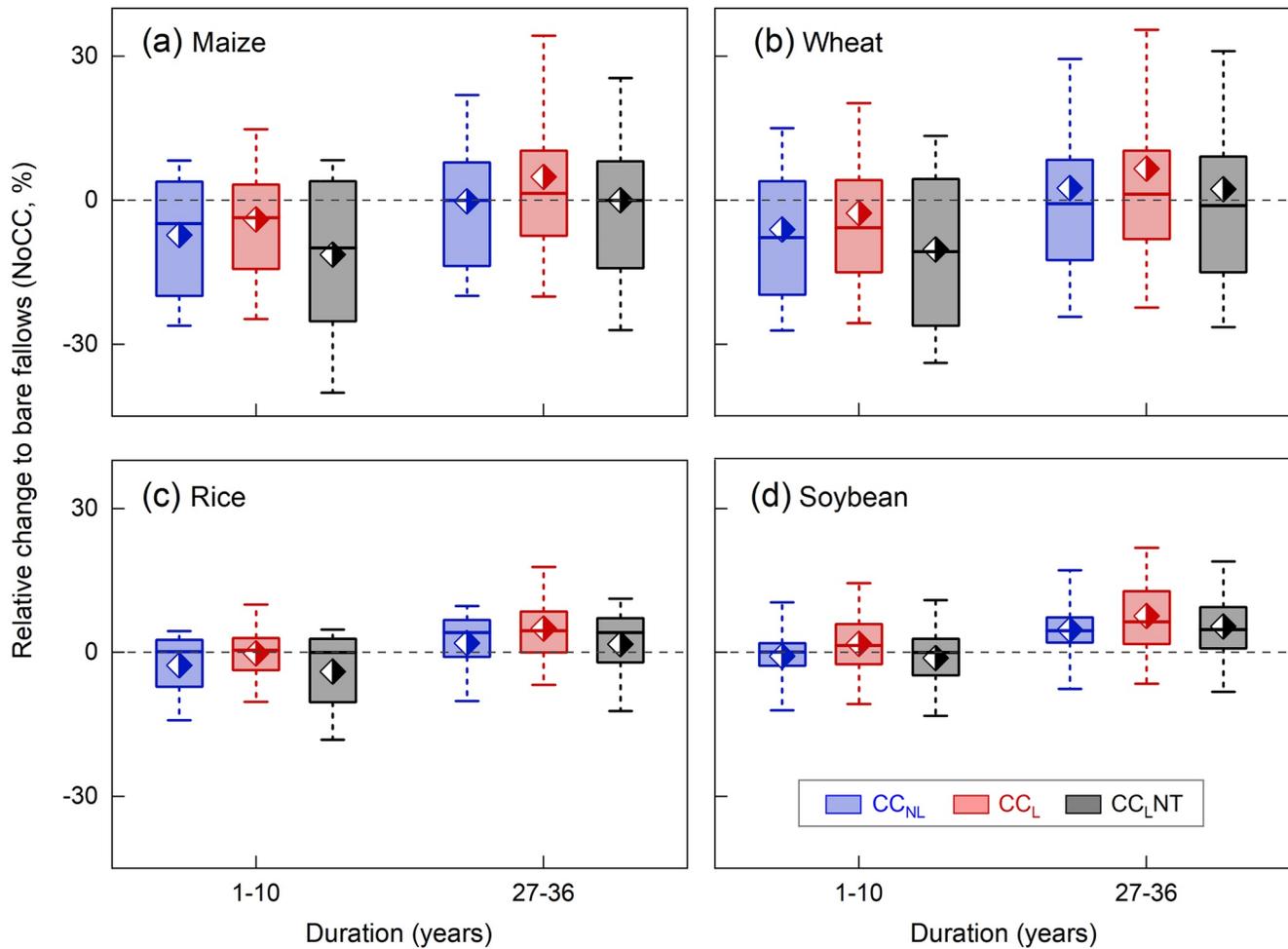
Figure 8. Box plots of the simulated crop-specific production in response to three cover crop managements (CC$_{NL}$: blue; CC$_L$: red; CC$_L$NT: black), compared to the control treatment with bare fallows (NoCC) in the first and last simulated decades under the CA$_{all}$ scenario: maize (a), wheat (b), rice (c), and soybean (d). Modeled crop-specific production at each grid cell was calculated as the area-weighted aggregated results in rain-fed and irrigated conditions. Box plots of yield relative changes (%) denote the 5th and 95th percentiles with whiskers, median and interquartile range with box lines, and mean with diamonds across all crop-specific grid cells (maize: 31,635; wheat: 27,126; rice: 21,598; soybean: 23,306).

crop types included, ~15.5% higher than the bare-fallow control treatment. In our model experiments, only the combined agricultural practice, that is, legume CCs and no tillage (CC$_L$NT), produced a mean SOC increase of 9.7% after a 36-year simulation (Figure 5), which is more comparable to but still below the findings in Jian et al. (2020). The discrepancy between the global simulation and site-level field experiments likely reflects their difference in the investigated geographical scales and land-use history, as well as to the diverse managements and methodologies among field studies (such as CC species, retained residue proportion, and implementation duration). Nevertheless, the potential of obtaining higher SOC stocks via cover crop management seems realistic, even though the exact magnitude of the effect remains unresolved.

In the global experiment, the annual SOC sequestration rate was modeled to be largest in the early years after introduction of CCs, and it then gradually declined over the course of the remaining simulation period (Figure 5), similar to published findings. Sommer and Bossio (2014) reported annual SOC stock changes in response to the improved agricultural practices approaching a maximum between the third and seventh year after adopting soil-conserving techniques and a subsequent decreasing trend for 15–20 years. A meta-analysis of tropical crop ecosystems also indicated reduced SOC sequestration rates (after an initial peak) to persist for 4–25 years until a new SOC equilibrium state was reached, but the duration was highly dependent on climates and soil types (Powlson et al., 2016). In our model experiments, at the end of 36-year simulation the continued trends indicate that a new steady state in soil C and N pools had not yet been achieved, which was similar to results in Porwollik

et al. (2022), who found no dynamic steady state after 50 years of simulation with the LPJmL model in response to planting herbaceous CCs on global cropland during fallow period.

In our study we attempted to quantify the contribution of CCs to enhancing soil C pools globally, which could also be interpreted as a climate change mitigation measure. After 36 years of implementation, using two herbaceous CCs was found to sequester ~0.01 Pg yr^{-1} soil carbon across the simulated 126×10^6 ha cropland (CA_{his} scenario, ~8% of current cropland areas worldwide; Table 2). If all agricultural lands were to adopt cover crop practices (CA_{all} scenario), the SOC sequestration potential could be as high as 0.11, 0.15, and 0.32 Pg C yr^{-1} (i.e., 0.40, 0.55, and 1.17 Pg CO_2 yr^{-1}) for non-legumes (CC_{NL}), legumes (CC_L), and the combined agricultural practice (CC_LNT), respectively, compensating for 8%–22% of annual direct GHG emissions from crops and livestock activities (5.3 Pg CO_2eq yr^{-1}; FAO, 2020), or equivalent to 10%–29% of GHG emissions from agricultural land use change (4.0 Pg CO_2eq yr^{-1}; FAO, 2020). Planting anywhere near 100% of global cropland with CCs is impractical for a number of reasons: a large share of agricultural area used for winter crops (Kaye & Quemada, 2017; Poeplau & Don, 2015), potential water limitations or too low winter temperature during off-season periods (Dabney et al., 2001), and insufficient growing windows for CCs in multi-cropping systems in the tropics (Hu et al., 2018; Zhu et al., 2012). Nevertheless, these estimates from our simulations do provide an upper bound for the amount of atmospheric carbon that might be sequestered through cover crop cultivation. Under the more realistic adoption scenario of CA_{pot} (590×10^6 ha, ~37% of current cropland areas; Table 2), carbon taken up in response to individual cover crop practices (0.15 and 0.22 Pg CO_2 yr^{-1} for CC_{NL} and CC_L, respectively) and the combined conservation management (CC_LNT; 0.46 Pg CO_2 yr^{-1}) could approximately offset 3%–9% of direct yearly GHG emissions from crops and livestock activities. However, additional N inputs to the soil from CCs could also potentially offset the CO_2 mitigation effect on the field scale as these would lead to increased N_2O emissions (Lugato et al., 2018; Quemada et al., 2020). Whether such a trade-off between soil carbon and nitrogen GHG fluxes due to cover cropping would emerge at the global level was not considered in this study and thus needs to be quantified in future modeling work.

4.2. N Leaching

Both model and field experiments showed that N leaching from cropland ecosystems was strongly associated with N management: applying chemical fertilizer resulted in higher hydrological N loss compared with the unfertilized treatments (Figure 4 and Figure S3 in Supporting Information S1), likely a consequence of the enhanced size of the nitrate pool. However, several disagreements between simulated and measured N leaching were found for some field trial locations despite similar N fertilizer inputs (Figure 2 and Figure S3 in Supporting Information S1), indicating that other factors, such as soil texture type, climate condition, or throughflow, are at play as well. For example, two of the field experiments included in our analysis sites showed a decreasing trend in total N leaching (mineral plus organic) from coarse-, medium-, to fine-textured soils (Aronsson et al., 2011; Lemola & Turtola, 2000). When testing our simulation setup at these two locations, the reported soil texture effect was not captured well by the model (not shown), suggesting that the N leaching representation in LPJ-GUESS should be further improved. Moreover, compared with observations, the overall smaller reduction in N leaching in response to the simulated CCs (Figure 2) might be partially attributed to the underestimated biomass of CCs (Figures 3 and 4), which would also underestimate plant N demand and soil N uptake. In addition, since the model cannot simulate two plants growing at the same time (see Section 2.3.1), the total length of the undersown-CC growing period in our simulations was approximately 1–2 months shorter than the field trials across all northern European sites (Table S3 in Supporting Information S1), which further limited cover crop capacity for uptake of excess N remaining in the soil column in the model.

Compared with the bare-fallow setup, mean decreases of 41% and 34% in N leaching were simulated across the globe in response to the experiment with non-legume (CC_{NL}) and legume cover crops (CC_L), respectively (Table 2), close to the lower end of the wide reported reduction range between 30% and 70% in the literature (Abdalla et al., 2019; Nouri et al., 2022; Quemada et al., 2013; Thapa et al., 2018; Tonitto et al., 2006). The reduction in N leaching due to CCs partially reflects the decreases in leachate volume and soil reactive N concentration because of enhanced water and N uptake by CCs during their growth (Thapa et al., 2018). This process may also underlie the smaller decreases in N leaching under N-fixing CCs compared with non-legumes for both field measurements (Abdalla et al., 2019; Nouri et al., 2022) and model simulations (CC_{NL} vs. CC_L). Where biological N fixation is the dominant N source for leguminous plants, it diminishes the capacity for mineral N uptake from

soils (Fontaine et al., 2022). Moreover, including the no-till technique in cover cropping in our simulations had the potential to further mitigate N leaching (41% in CC_LNT vs. 34% in CC_L; Table 2) mainly due to the reduced net N mineralization rates (Figure S7 in Supporting Information S1). This is in line with the findings from a meta-analysis by Thapa et al. (2018) and a recent modeling study by Porwollik et al. (2022).

Globally, the largest percent decreases in N leaching due to CCs were modeled in regions with relatively little N fertilizer use (such as Russia and large parts of Africa; Figure 6 and Figure S1 in Supporting Information S1), where soil reactive N pools were small. Results from a 6-year field experiment implemented by Wittwer et al. (2017) also showed that the effectiveness of CCs in reducing N leaching decreased with management intensity (e.g., tillage regimes and fertilizer application rates). This effect underlies discrepancies at some national borders, such as Indonesia and Papua New Guinea (Figure 6), countries with similar climates but with contrasting fertilizer applications (Figure S1 in Supporting Information S1). Likewise, in some arid and semi-arid regions, as well as temperature-limited areas in the high latitudes (e.g., Canada) a slight decrease of N leaching in response to cover cropping systems was found, as poor growth conditions constrained the CC capacity for soil N uptake. In addition, the rapid turnover rate of SOM pools driven by warm and moist climate (Olin, Lindeskog, et al., 2015), together with abundant precipitation may increase N leaching with cover crop practices in the humid tropics (Figure 6) as a result of high biomass of N returned to soils (Figure S6 in Supporting Information S1) and enhanced throughflow (Porwollik et al., 2022).

4.3. Crop Yields

Accounting for the impacts of management practices, particularly regarding water and N limitations to crop growth in LPJ-GUESS, resulted in a good agreement between simulated and observed crop yields across different field trials despite some outliers in rice and wheat systems (Figures S3 and S8 in Supporting Information S1). For both modeling and field-based experiments, yields in the main crops following non-legume CCs declined, although the overall difference from fallow controls (NoCC) was small (Figure 2b). The difference between periods of soil N mineralization and high N demand of main crops (Marcillo & Miguez, 2017), and enhanced soil N immobilization shortly after the planting of non-legume CCs (Abdalla et al., 2019; Erenstein, 2003) may contribute to the declines in yields of the main crops in the field experiments. In comparison, N-fixing CCs with relatively low C:N ratios are expected to stimulate soil N release during their decomposition, enhancing plant-available N in soils (Li et al., 2020; Quemada et al., 2013; Thapa et al., 2018). This was in line with our model findings, wherein legume CCs generally resulted in higher net N mineralization rates than non-legumes (Figure S7 in Supporting Information S1) and thus increased the productivity of the main crops in some cases (Figure 2b). However, it should be noted that these CC effects were highly dependent on cropping systems, with little impact found on productivity of soybeans (Table S5 in Supporting Information S1). This is likely due to their N fixation capacity, which diminished the N competition between CCs and soybeans in both field trials and model simulations.

Our modeled global mean yield losses due to CCs in the first decade of the simulations (−3% for CC_L and −6% for CC_{NL}; CA_{all} scenario in Table 2) compared well with a recent meta-analysis by Garba et al. (2022), who reported a mean crop production change of −4.9% and −10.1% for legume and non-legume CCs, respectively, after 2–17 years of management. Main-crop yield reduction under cover cropping systems likely reflected (a) the indirect competition for water and nutrients between CCs and subsequent main crops (Valkama et al., 2015), and (b) the time that soil SOM pools need to adjust to management shifts (Figure 5 and Figure S7 in Supporting Information S1). Garba et al. (2022) also pointed out that cover cropping systems under the no-till practice resulted in lower main-crop yields compared with conventional tillage, in line with our model findings in terms of total crop production worldwide (CC_LNT vs. CC_L; Table 2). However, at least in our simulations, these negative yield effects induced by conservation tillage may be mitigated over the course of the simulation (Table 2) because of the gradual stabilization of soil C and N pools over time (Figure 5 and Figure S7 in Supporting Information S1). A similar finding from a meta-analysis by Pittelkow et al. (2015) indicated that yield benefits, globally, in cereal- and legume-based cropping systems may be attained after 10+ years of conversion from conventional tillage to no-till management.

N fertilizer application was found to be another factor that influenced the effectiveness of CCs on subsequent crop yields for both site-level (Figure S3 in Supporting Information S1) and large-scale simulations (Figure 7). The smallest impacts on main-crop production were found for well-fertilized cover cropping systems, consistent with

previous field-based reviews (Daryanto et al., 2018; Marcillo & Miguez, 2017; Quemada et al., 2013; Tonitto et al., 2006; Zhao et al., 2022), since enhanced soil mineral N pools driven by fertilization reduce the N competition between CCs and main crops. This can explain the small yield penalty (or benefit) from cover cropping in soybean (Figures 7 and 8), which is a nitrogen fixer and experiences less N stress during the growing season compared with cereal crops. Likewise, the spatial variability regarding CC impacts on rice production was also much smaller than simulated maize and wheat CFTs (Figures 7 and 8), primarily because rice in our simulations was mostly irrigated (Figure 7), which reduced water limitation on crop growth caused by CCs in rice-producing areas. Furthermore, the broadly negative impacts of CCs on simulated yields in northern temperate climatic regions (Figure 7) can be attributed to the slow decomposition of SOM in response to low temperature, where the N retained in the SOM is released evenly throughout the year and not easily available for main crop uptake after CC growth (Olin, Lindeskog, et al., 2015). In contrast and as discussed above, plant materials from CCs in the humid tropics are expected to rapidly decompose due to the fast turnover rate, continuously releasing reactive N for plant uptake in the next cropping season and therefore enhancing main-crop productions. This contrasting spatial difference in yield changes between temperate and tropical climates supports a meta-analysis finding that cultivating CCs during bare-fallow period, on average, has a risk to reduce main-crop productivity by ∼12% in temperate agricultural soils while gaining ∼15% of yield benefits in the tropics (Garba et al., 2022).

4.4. Modeling Limitations and Implications

A detailed evaluation of modeling CC impacts on cropland worldwide remains a challenge due to various cover crop species, farming rotation systems, and managements in the field trials. We mainly examined the model performance via categorizing herbaceous CCs as non-legume and legume functional types, with site-specific management practices considered (Tables S2–S4 in Supporting Information S1). Although the current C-N version of LPJ-GUESS can reproduce the observed responses of ecosystem service indicators to CC cultivation, the magnitude of these changes did not always match experimental measurements (Figure 2, Table S5 in Supporting Information S1). This likely reflects the differences between highly controlled field conditions and model's representation of management history, initial SOM levels, cropping system management, and the C-N allocation scheme in CCs. In addition, important processes that determine CC impacts in the field experiments—such as occurrence of weeds (Mazzoncini et al., 2011), intercropping (Valkama et al., 2015), termination methods of CCs (Bloszies et al., 2022); erosion (Daryanto et al., 2018), and soil structural modification via grass roots (Nouri et al., 2022)—have not been accounted for in the model.

To compare model outputs with observations, as introduced in Section 2.4, we standardized the measured SOC from the original depth to the modeled depth of 150 cm using an empirical depth distribution function. There are large uncertainties associated with these extrapolated SOC stocks due to the varying management effects on soil C pools with depth. For example, a global meta-analyses is found SOC benefits of no-till farming to be statistically significant in the topsoil (0–15 cm) and decline with soil depth (Haddaway et al., 2017). Scaling SOC stocks with a simple extrapolation function cannot reflect the observed variability in the field and thus our approach by necessity is a simplified one.

Legume CCs are usually identified as a promising strategy to substitute chemical N fertilizer in agricultural productions due to their high N fixation rates (Herridge et al., 2022; Peoples et al., 2021). Our modeled N fixed by natural C_3 grass (a surrogate for white clover; see Text S1 in Supporting Information S1) during main-crop off-season periods are 30–70 kg N ha^{-1} yr^{-1} in warm and moist regions (36-year average; Figure S6 in Supporting Information S1), which are lower than the reported range of 49–154 kg N ha^{-1} yr^{-1} but these latter estimates were for the entire year (Anglade et al., 2015; Burchill et al., 2014; Ledgard et al., 2001). Nonetheless, in our simulations employing legume CCs results in higher yield benefits in the humid tropics compared with non-legumes (Figure 7 and Figure S5 in Supporting Information S1). As described in Section 2.1, one main growing season within a year is modeled in LPJ-GUESS, total agricultural production achieved by multi-cropping systems in the tropics are not yet captured. As a consequence, the N fixation rate and biomass in legume CCs may be too high since we overestimate the length of the bare-fallow period for cover crop cultivation (Porwollik et al., 2022). Compared to controls with no CCs, such an overestimation would then be possibly reflected in high SOC sequestration rates and yield benefits in tropical climates.

The inclusion of the no-till technique in cover cropping is an effective practice under CA systems for mitigating climate change (Blanco-Canqui et al., 2015). This combined strategy in our study is also modeled as a win-win

management option in terms of enhancing SOC stocks while reducing N leaching rates, despite the accompanying ∼8% of decrease in total crop production when integrated over global cropland for the first simulated decade (Table 2). It should be noted that assessing the effects of no-till management on cropland N leaching remains uncertain. Some studies reported that conservation tillage can slightly reduce this hydrological N loss because of the diminished net N mineralization rates (Porwollik et al., 2022; Salahin et al., 2021; Thapa et al., 2018). However, other studies found enhanced nitrate leaching in the reduced-tillage soils compared with conventional tillage systems, mainly due to the enhanced water drainage caused by greater abundance of macropores (preferential flow channels) and better soil infiltrability (Daryanto et al., 2017). It remains unknown which of these two processes played a more important role in the field trials, but the modifications of soil structural and hydraulic properties in response to tillage are not included in the version of the LPJ-GUESS used in this study.

Rather than planting herbaceous CCs, it is more common to use legume crops (e.g., faba bean and field peas) as "green manure" in some temperate regions (Andersen et al., 2020; Rinnofner et al., 2008). These grain legumes are usually intercropped with other cash crops, and incorporated to soils at full bloom stage to maximize N fixation rates while minimizing soil water depletion (Denton et al., 2017; Williams et al., 2014). To better represent region-specific cover crop practices, the implementation of N-fixing grain legumes as intercrops, together with multi-cropping systems within a year (see discussion above), remains to be taken into account in future model work.

5. Conclusions

In this study we developed a new C_3 grass functional type with biological N fixation in LPJ-GUESS to better account for legume CC effects on global crop ecosystems. The simulated C-N variables and main-crop productions in response to two herbaceous cover crop types (i.e., non-legumes and legumes) were widely evaluated against measured data from site level to global. Our model estimates demonstrated that crop ecosystems implemented in LPJ-GUESS realistically responded to non-legume and legume cover cropping under a range of water and N managements, and resulted in comparable C-N variables with observations, particularly for cropland SOC stocks.

When integrated over global croplands, our long-term simulations revealed that the impacts of CCs on agricultural soils can be beneficial for environmental sustainability without compromising crop production, particularly for the integrated management practice with legume CCs and no-till technique included. This combined strategy was modeled to achieve an annual SOC sequestration rate of 0.32 Mg C ha^{-1} yr^{-1} and to reduce N leaching by 41% (36-year average), also with a yield increase of 2% in the last simulated decade. The influence of CCs on crop production was strongly associated with main crop types and N fertilizer inputs, with small yield changes found in soybean systems and highly fertilized agricultural soils. Processes missing in the model, such as weeds, within-year multi-cropping systems, and cover crop management, may have biased our estimates of CC impacts on cropland globally.

The dynamic process of N fixation for grass CCs in LPJ-GUESS provides an opportunity to overall assess atmospheric carbon and nitrogen flows to agricultural lands during fallow periods, and thus is relevant for the estimates of global terrestrial C-N fluxes and pools under present-day and future climate, including how CO_2 uptake versus N_2O emissions might interplay. It can also help to predict the possibility of substituting synthetic fertilizer with N-fixing green manure in global crop ecosystems, with various management strategies and climate conditions considered.

Conflict of Interest

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

Data Availability Statement

Global historical climate data of GSWP3-W5E5 are available at https://doi.org/10.48364/ISIMIP.342217 (Lange et al., 2021). The monthly climate forcing data set of CRUJRA can be downloaded at https://data.ceda.ac.uk/badc/cru/data/cru_jra/cru_jra_2.1 (Harris et al., 2020; Kobayashi et al., 2015). National yield statistics of four crop types presented in this paper are from http://www.fao.org/faostat/en/#data (FAOSTAT, 2023). The site-level observations collected from the existing literature, together with large-scale model inputs and outputs as shown

in the figures of this study, can be publicly accessed through the Zenodo repository at https://doi.org/10.5281/zenodo.7646911 (Ma et al., 2023).

References

Abdalla, M., Hastings, A., Cheng, K., Yue, Q., Chadwick, D., Espenberg, M., et al. (2019). A critical review of the impacts of cover crops on nitrogen leaching, net greenhouse gas balance and crop productivity. *Global Change Biology*, *25*(8), 2530–2543. https://doi.org/10.1111/gcb.14644

Andersen, B. J., Samarappuli, D. P., Wick, A., & Berti, M. T. (2020). Faba bean and pea can provide late-fall forage grazing without affecting maize yield the following season. *Agronomy*, *10*(1), 80. https://doi.org/10.3390/agronomy10010080

Anglade, J., Billen, G., & Garnier, J. (2015). Relationships for estimating N2 fixation in legumes: Incidence for N balance of legume-based cropping systems in Europe. *Ecosphere*, *6*(3), 1–24. https://doi.org/10.1890/ES14-00353.1

Arneth, A., Olsson, L., Cowie, A., Erb, K. H., Hurlbert, M., Kurz, W. A., et al. (2021). Restoring degraded lands. *Annual Review of Environment and Resources*, *46*(1), 569–599. https://doi.org/10.1146/annurev-environ-012320-054809

Aronsson, H., Stenberg, M., & Ulén, B. (2011). Leaching of N, P and glyphosate from two soils after herbicide treatment and incorporation of a ryegrass catch crop. *Soil Use & Management*, *27*(1), 54–68. https://doi.org/10.1111/j.1475-2743.2010.00311.x

Blanco-Canqui, H., Shaver, T. M., Lindquist, J. L., Shapiro, C. A., Elmore, R. W., Francis, C. A., & Hergert, G. W. (2015). Cover crops and ecosystem services: Insights from studies in temperate soils. *Agronomy Journal*, *107*(6), 2449–2474. https://doi.org/10.2134/agronj15.0086

Bloszies, S. A., Reberg-Horton, S. C., Heitman, J. L., Woodley, A. L., Grossman, J. M., & Hu, S. (2022). Legume cover crop type and termination method effects on labile soil carbon and nitrogen and aggregation. *Agronomy Journal*, *114*(3), 1817–1832. https://doi.org/10.1002/agj2.21022

Burchill, W., James, E. K., Li, D., Lanigan, G. J., Williams, M., Iannetta, P. P. M., & Humphreys, J. (2014). Comparisons of biological nitrogen fixation in association with white clover (Trifolium repens L.) under four fertiliser nitrogen inputs as measured using two 15N techniques. *Plant and Soil*, *385*(1–2), 287–302. https://doi.org/10.1007/s11104-014-2199-1

Chatskikh, D., Hansen, S., Olesen, J. E., & Petersen, B. M. (2009). A simplified modelling approach for quantifying tillage effects on soil carbon stocks. *European Journal of Soil Science*, *60*(6), 924–934. https://doi.org/10.1111/j.1365-2389.2009.01185.x

Chatterjee, N., Archontoulis, S. V., Bastidas, A., Proctor, C. A., Elmore, R. W., & Basche, A. D. (2020). Simulating winter rye cover crop production under alternative management in a corn-soybean rotation. *Agronomy Journal*, *112*(6), 4648–4665. https://doi.org/10.1002/agj2.20377

Ciampitti, I. A., & Salvagiotti, F. (2018). New insights into soybean biological nitrogen fixation. *Agronomy Journal*, *110*(4), 1185–1196. https://doi.org/10.2134/agronj2017.06.0348

Constantin, J., Mary, B., Laurent, F., Aubrion, G., Fontaine, A., Kerveillant, P., & Beaudoin, N. (2010). Effects of catch crops, no till and reduced nitrogen fertilization on nitrogen leaching and balance in three long-term experiments. *Agriculture, Ecosystems & Environment*, *135*(4), 268–278. https://doi.org/10.1016/j.agee.2009.10.005

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., et al. (2020). WFDE5: Bias-adjusted ERA5 reanalysis data for impact studies. *Earth System Science Data*, *12*(3), 2097–2120. https://doi.org/10.5194/essd-12-2097-2020

Dabney, S. M., Delgado, J. A., & Reeves, D. W. (2001). Using winter cover crops to improve soil and water quality. *Communications in Soil Science and Plant Analysis*, *32*(7–8), 1221–1250. https://doi.org/10.1081/CSS-100104110

Daryanto, S., Fu, B., Wang, L., Jacinthe, P. A., & Zhao, W. (2018). Quantitative synthesis on the ecosystem services of cover crops. *Earth-Science Reviews*, *185*(June), 357–373. https://doi.org/10.1016/j.earscirev.2018.06.013

Daryanto, S., Wang, L., & Jacinthe, P. A. (2017). Impacts of no-tillage management on nitrate loss from corn, soybean and wheat cultivation: A meta-analysis. *Scientific Reports*, *7*(1), 1–9. https://doi.org/10.1038/s41598-017-12383-7

Denton, M. D., Phillips, L. A., Peoples, M. B., Pearce, D. J., Swan, A. D., Mele, P. M., & Brockwell, J. (2017). Legume inoculant application methods: Effects on nodulation patterns, nitrogen fixation, crop growth and yield in narrow-leaf lupin and faba bean. *Plant and Soil*, *419*(1–2), 25–39. https://doi.org/10.1007/s11104-017-3317-7

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., & Hanasaki, N. (2006). GSWP-2: Multimodel analysis and implications for our perception of the land surface. *Bulletin of the American Meteorological Society*, *87*(10), 1381–1397. https://doi.org/10.1175/BAMS-87-10-1381

Don, A., Schumacher, J., & Freibauer, A. (2011). Impact of tropical land-use change on soil organic carbon stocks—A meta-analysis. *Global Change Biology*, *17*(4), 1658–1670. https://doi.org/10.1111/j.1365-2486.2010.02336.x

Duval, M. E., Galantini, J. A., Capurro, J. E., & Martinez, J. M. (2016). Winter cover crops in soybean monoculture: Effects on soil organic carbon and its fractions. *Soil and Tillage Research*, *161*, 95–105. https://doi.org/10.1016/j.still.2016.04.006

Elliott, J., Müller, C., Deryng, D., Chryssanthacopoulos, J., Boote, K. J., Büchner, M., et al. (2015). The global gridded crop model intercomparison: Data and modeling protocols for phase 1 (v1.0). *Geoscientific Model Development*, *8*(2), 261–277. https://doi.org/10.5194/gmd-8-261-2015

Erenstein, O. (2003). Smallholder conservation farming in the tropics and sub-tropics: A guide to the development and dissemination of mulching with crop residues and cover crops. *Agriculture, Ecosystems & Environment*, *100*(1–3), 17–37. https://doi.org/10.1016/S0167-8809(03)00150-6

FAO. (2020). Emissions due to agriculture. Global, regional and country trends 2000-2018. FAOSTAT Analytical Brief Series No 18. Retrieved from https://www.fao.org/3/cb3808en/cb3808en.pdf

FAOSTAT (2023). Production/Crops and livestock products. [Dataset]. FAOSTAT. Retrieved from https://www.fao.org/faostat/en/#data

Fontaine, D., Kristensen, R. K., Rasmussen, J., & Eriksen, J. (2022). Data on growth, uptake and N2 fixation of grass-clover leys fertilized with mineral N fertilizer and cattle slurry. *Data in Brief*, *42*, 107998. https://doi.org/10.1016/j.dib.2022.107998

Garba, I. I., Bell, L. W., & Williams, A. (2022). Cover crop legacy impacts on soil water and nitrogen dynamics, and on subsequent crop yields in dryland: A meta-analysis. *Agronomy for Sustainable Development*, *42*(3), 34. https://doi.org/10.1007/s13593-022-00760-0

Global Soil Data Task. (2014). *Global soil data products CD-ROM contents*. IGBP-DIS. https://doi.org/10.3334/ORNLDAAC/565

Haddaway, N. R., Hedlund, K., Jackson, L. E., Kätterer, T., Lugato, E., Thomsen, I. K., et al. (2017). How does tillage intensity affect soil organic carbon? A systematic review. *Environmental Evidence*, *6*(30), 30. https://doi.org/10.1186/s13750-017-0108-9

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, *7*(1), 109. https://doi.org/10.1038/s41597-020-0453-3

Herridge, D. F., Giller, K. E., Jensen, E. S., & Peoples, M. B. (2022). Quantifying country-to-global scale nitrogen fixation for grain legumes II. Coefficients, templates and estimates for soybean, groundnut and pulses. *Plant and Soil*, *474*(1–2), 1–15. https://doi.org/10.1007/s11104-021-05166-7

Hu, A., Tang, T., & Liu, Q. (2018). Nitrogen use efficiency in different rice-based rotations in southern China. *Nutrient Cycling in Agroecosystems*, *112*(1), 75–86. https://doi.org/10.1007/s10705-018-9930-x

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., et al. (2020). Harmonization of global land use change and management for the period 2015-2300 (LUH2) for CMIP6. *Geoscientific Model Development*, *13*(11), 5425–5464. https://doi.org/10.5194/gmd-13-5425-2020

Jang, W. S., Neff, J. C., Im, Y., Doro, L., & Herrick, J. E. (2021). The hidden costs of land degradation in US maize agriculture. *Earth's Future*, *9*(2), e2020EF001641. https://doi.org/10.1029/2020EF001641

Jian, J., Du, X., Reiter, M. S., & Stewart, R. D. (2020). A meta-analysis of global cropland soil carbon changes due to cover cropping. *Soil Biology and Biochemistry*, *143*, 107735. https://doi.org/10.1016/j.soilbio.2020.107735

Jobbágy, E. G., & Jackson, R. B. (2000). The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecological Applications*, *10*(2), 423–436. https://doi.org/10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2

Kaschuk, G., Hungria, M., Leffelaar, P. A., Giller, K. E., & Kuyper, T. W. (2010). Differences in photosynthetic behaviour and leaf senescence of soybean (Glycine max [L.] Merrill) dependent on N2 fixation or nitrate supply. *Plant Biology*, *12*(1), 60–69. https://doi.org/10.1111/j.1438-8677.2009.00211.x

Kaschuk, G., Kuyper, T. W., Leffelaar, P. A., Hungria, M., & Giller, K. E. (2009). Are the rates of photosynthesis stimulated by the carbon sink strength of rhizobial and arbuscular mycorrhizal symbioses? *Soil Biology and Biochemistry*, *41*(6), 1233–1244. https://doi.org/10.1016/j.soilbio.2009.03.005

Kaspar, T. C., Jaynes, D. B., Parkin, T. B., Moorman, T. B., & Singer, J. W. (2012). Effectiveness of oat and rye cover crops in reducing nitrate losses in drainage water. *Agricultural Water Management*, *110*(3), 25–33. https://doi.org/10.1016/j.agwat.2012.03.010

Kaye, J. P., & Quemada, M. (2017). Using cover crops to mitigate and adapt to climate change. A review. *Agronomy for Sustainable Development*, *37*(1), 4. https://doi.org/10.1007/s13593-016-0410-x

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan*, *93*(1), 5–48. https://doi.org/10.2151/jmsj.2015-001

Kuo, S., Sainju, U. M., & Jellum, E. J. (1997). Winter cover crop effects on soil organic carbon and carbohydrate in soil. *Soil Science Society of America Journal*, *61*(1), 145–152. https://doi.org/10.2136/sssaj1997.03615995006100010022x

Lal, R. (2004). Soil carbon sequestration impacts on global climate change and food security. *Science*, *304*(5677), 1623–1627. https://doi.org/10.1126/science.1097396

Lange, S. (2019). Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development*, *12*, 3055–3070. https://doi.org/10.5194/gmd-2019-36

Lange, S., Menz, C., Gleixner, S., Cucchi, M., Weedon, G. P., Amici, A., et al. (2021). WFDE5 over land merged with ERA5 over the ocean (W5E5 v2.0). [Dataset]. ISIMIP. https://doi.org/10.48364/ISIMIP.342217

Ledgard, S. F., Sprosen, M. S., Penno, J. W., & Rajendram, G. S. (2001). Nitrogen fixation by white clover in pastures grazed by dairy cows: Temporal variation and effects of nitrogen fertilization. *Plant and Soil*, *229*(2), 177–187. https://doi.org/10.1023/A:1004833804002

Lemola, R., Turtola, E., & Eriksson, C. (2000). Undersowing Italian ryegrass diminishes nitrogen leaching from spring barley. *Agricultural and Food Science in Finland*, *9*(3), 201–215. https://doi.org/10.23986/afsci.5661

Li, F., Sørensen, P., Li, X., & Olesen, J. E. (2020). Carbon and nitrogen mineralization differ between incorporated shoots and roots of legume versus non-legume based cover crops. *Plant and Soil*, *446*(1–2), 243–257. https://doi.org/10.1007/s11104-019-04358-6

Lin, B. L., Sakoda, A., Shibasaki, R., & Suzuki, M. (2001). A modelling approach to global nitrate leaching caused by anthropogenic fertilisation. *Water Research*, *35*(8), 1961–1968. https://doi.org/10.1016/S0043-1354(00)00484-X

Lindeskog, M., Arneth, A., Bondeau, A., Waha, K., Seaquist, J., Olin, S., & Smith, B. (2013). Implications of accounting for land use in simulations of ecosystem carbon cycling in Africa. *Earth System Dynamics*, *4*(2), 385–407. https://doi.org/10.5194/esd-4-385-2013

Liu, J., You, L., Amini, M., Obersteiner, M., Herrero, M., Zehnder, A. J. B., & Yang, H. (2010). A high-resolution assessment on global nitrogen flows in cropland. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(17), 8035–8040. https://doi.org/10.1073/pnas.0913658107

Liu, L., Xu, W., Lu, X., Zhong, B., & Guo, Y. (2022). Exploring global changes in agricultural ammonia emissions and their contribution to nitrogen deposition since 1980. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(14), e2121998119. https://doi.org/10.1073/pnas.2121998119

Liu, Q., Liu, B., Zhang, Y., Hu, T., Lin, Z., Liu, G., et al. (2019). Biochar application as a tool to decrease soil nitrogen losses (NH3 volatilization, N2O emissions, and N leaching) from croplands: Options and mitigation strength in a global perspective. *Global Change Biology*, *25*(6), 2077–2093. https://doi.org/10.1111/gcb.14613

Lu, C., & Tian, H. (2017). Global nitrogen and phosphorus fertilizer use for agriculture production in the past half century: Shifted hot spots and nutrient imbalance. *Earth System Science Data*, *9*(1), 181–192. https://doi.org/10.5194/essd-9-181-2017

Lugato, E., Leip, A., & Jones, A. (2018). Mitigation potential of soil carbon management overestimated by neglecting N2O emissions. *Nature Climate Change*, *8*(3), 219–223. https://doi.org/10.1038/s41558-018-0087-z

Luo, Z., Wang, E., & Sun, O. J. (2010). Can no-tillage stimulate carbon sequestration in agricultural soils? A meta-analysis of paired experiments. *Agriculture, Ecosystems & Environment*, *139*(1–2), 224–231. https://doi.org/10.1016/j.agee.2010.08.006

Lutz, F., Del Grosso, S., Ogle, S., Williams, S., Minoli, S., Rolinski, S., et al. (2020). The importance of management information and soil moisture representation for simulating tillage effects on N2O emissions in LPJmL5.0-tillage. *Geoscientific Model Development*, *13*(9), 3905–3923. https://doi.org/10.5194/gmd-13-3905-2020

Ma, J., Anthoni, P., Olin, S., Rabin, S. S., Bayer, A. D., Xia, L., & Arneth, A. (2023). Cover crops effect on global croplands modelled by LPJ-GUESS. [Dataset]. zenodo. https://doi.org/10.5281/zenodo.7646911

Ma, J., Olin, S., Anthoni, P., Rabin, S. S., Bayer, A. D., Nyawira, S. S., & Arneth, A. (2022). Modeling symbiotic biological nitrogen fixation in grain legumes globally with LPJ-GUESS (v4.0, r10285). *Geoscientific Model Development*, *15*(2), 815–839. https://doi.org/10.5194/gmd-15-815-2022

Ma, J., Rabin, S. S., Anthoni, P., Bayer, A. D., Nyawira, S. S., Olin, S., et al. (2022). Assessing the impacts of agricultural managements on soil carbon stocks, nitrogen loss and crop production - A modelling study in eastern Africa. *Biogeosciences*, *19*(8), 2145–2169. https://doi.org/10.5194/bg-19-2145-2022

Marcillo, G. S., & Miguez, F. E. (2017). Corn yield response to winter cover crops: An updated meta-analysis. *Journal of Soil and Water Conservation*, *72*(3), 226–239. https://doi.org/10.2489/jswc.72.3.226

Mazzoncini, M., Sapkota, T. B., Bàrberi, P., Antichi, D., & Risaliti, R. (2011). Long-term effect of tillage, nitrogen fertilization and cover crops on soil organic carbon and total nitrogen content. *Soil and Tillage Research*, *114*(2), 165–174. https://doi.org/10.1016/j.still.2011.05.001

McClelland, S. C., Paustian, K., & Schipanski, M. E. (2021). Management of cover crops in temperate climates influences soil organic carbon stocks: A meta-analysis. *Ecological Applications*, *31*(3), 1–19. https://doi.org/10.1002/eap.2278

McDermid, S. S., Mearns, L. O., & Ruane, A. C. (2017). Representing agriculture in Earth system models: Approaches and priorities for development. *Journal of Advances in Modeling Earth Systems*, *9*(5), 2230–2265. https://doi.org/10.1002/2016MS000749

Meinshausen, M., Nicholls, Z. R. J., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., et al. (2020). The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. *Geoscientific Model Development*, *13*(8), 3571–3605. https://doi.org/10.5194/gmd-13-3571-2020

Moss, B. (2008). Water pollution by agriculture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1491), 659–666. https://doi.org/10.1098/rstb.2007.2176

Nouri, A., Lukas, S., Singh, S., Singh, S., & Machado, S. (2022). When do cover crops reduce nitrate leaching? A global meta-analysis. *Global Change Biology*, *28*(15), 4736–4749. https://doi.org/10.1111/gcb.16269

Olin, S., Lindeskog, M., Pugh, T. A. M., Schurgers, G., Wårlind, D., Mishurov, M., et al. (2015). Soil carbon management in large-scale Earth system modelling: Implications for crop yields and nitrogen leaching. *Earth System Dynamics*, *6*(2), 745–768. https://doi.org/10.5194/esd-6-745-2015

Olin, S., Schurgers, G., Lindeskog, M., Wårlind, D., Smith, B., Bodin, P., et al. (2015). Modelling the response of yields and tissue C: N to changes in atmospheric CO2 and N management in the main wheat regions of Western Europe. *Biogeosciences*, *12*(8), 2489–2515. https://doi.org/10.5194/bg-12-2489-2015

Parton, W. J., Scurlock, J. M. O., Ojima, D. S., Gilmanov, T. G., Scholes, R. J., Schimel, D. S., et al. (1993). Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide. *Global Biogeochemical Cycles*, *7*(4), 785–809. https://doi.org/10.1029/93GB02042

Penning de Vries, F. W. T., Jansen, D. M., ten Berge, H. F. M., & Bakema, A. (1989). *Simulation of ecophysiological processes of growth in several annual crops*. Centre for Agricultural Publishing and Documentation (Pudoc).

Peoples, M. B., Giller, K. E., Jensen, E. S., & Herridge, D. F. (2021). Quantifying country-to-global scale nitrogen fixation for grain legumes: I. Reliance on nitrogen fixation of soybean, groundnut and pulses. *Plant and Soil*, *469*(1–2), 1–14. https://doi.org/10.1007/s11104-021-05167-6

Pittelkow, C. M., Linquist, B. A., Lundy, M. E., Liang, X., van Groenigen, K. J., Lee, J., et al. (2015). When does no-till yield more? A global meta-analysis. *Field Crops Research*, *183*, 156–168. https://doi.org/10.1016/j.fcr.2015.07.020

Poeplau, C., & Don, A. (2015). Carbon sequestration in agricultural soils via cultivation of cover crops - a meta-analysis. *Agriculture, Ecosystems & Environment*, *200*, 33–41. https://doi.org/10.1016/j.agee.2014.10.024

Poeplau, C., Don, A., Vesterdal, L., Leifeld, J., Van Wesemael, B., Schumacher, J., & Gensior, A. (2011). Temporal dynamics of soil organic carbon after land-use change in the temperate zone - Carbon response functions as a model approach. *Global Change Biology*, *17*(7), 2415–2427. https://doi.org/10.1111/j.1365-2486.2011.02408.x

Pongratz, J., Dolman, H., Don, A., Erb, K. H., Fuchs, R., Herold, M., et al. (2018). Models meet data: Challenges and opportunities in implementing land management in Earth system models. *Global Change Biology*, *24*(4), 1470–1487. https://doi.org/10.1111/gcb.13988

Portmann, F. T., Siebert, S., & Döll, P. (2010). MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, *24*(1), 1–24. https://doi.org/10.1029/2008gb003435

Porwollik, V., Rolinski, S., Heinke, J., & Müller, C. (2019). Generating a rule-based global gridded tillage dataset. *Earth System Science Data*, *11*(2), 823–843. https://doi.org/10.5194/essd-11-823-2019

Porwollik, V., Rolinski, S., Heinke, J., von Bloh, W., Schaphoff, S., & Müller, C. (2022). The role of cover crops for cropland soil carbon, nitrogen leaching, and agricultural yields – A global simulation study with LPJmL (V. 5.0-tillage-cc). *Biogeosciences*, *19*(3), 957–977. https://doi.org/10.5194/bg-19-957-2022

Powlson, D. S., Stirling, C. M., Jat, M. L., Gerard, B. G., Palm, C. A., Sanchez, P. A., & Cassman, K. G. (2014). Limited potential of no-till agriculture for climate change mitigation. *Nature Climate Change*, *4*(8), 678–683. https://doi.org/10.1038/nclimate2292

Powlson, D. S., Stirling, C. M., Thierfelder, C., White, R. P., & Jat, M. L. (2016). Does conservation agriculture deliver climate change mitigation through soil carbon sequestration in tropical agro-ecosystems? *Agriculture, Ecosystems & Environment*, *220*, 164–174. https://doi.org/10.1016/j.agee.2016.01.005

Pugh, T. A. M., Arneth, A., Olin, S., Ahlström, A., Bayer, A. D., Klein Goldewijk, K., et al. (2015). Simulated carbon emissions from land-use change are substantially enhanced by accounting for agricultural management. *Environmental Research Letters*, *10*(12), 124008. https://doi.org/10.1088/1748-9326/10/12/124008

Qin, Z., Guan, K., Zhou, W., Peng, B., Tang, J., Jin, Z., et al. (2023). Assessing long-term impacts of cover crops on soil organic carbon in the central U. S. Midwestern agroecosystems. *Global Change Biology*, *29*(9), 2572–2590. https://doi.org/10.1111/gcb.16632

Quemada, M., Baranski, M., Nobel-de Lange, M. N. J., Vallejo, A., & Cooper, J. M. (2013). Meta-analysis of strategies to control nitrate leaching in irrigated agricultural systems and their effects on crop yield. *Agriculture, Ecosystems & Environment*, *174*, 1–10. https://doi.org/10.1016/j.agee.2013.04.018

Quemada, M., Lassaletta, L., Leip, A., Jones, A., & Lugato, E. (2020). Integrated management for sustainable cropping systems: Looking beyond the greenhouse balance at the field scale. *Global Change Biology*, *26*(4), 2584–2598. https://doi.org/10.1111/gcb.14989

Rasse, D. P., Rumpel, C., & Dignac, M.-F. (2005). Is soil carbon mostly root carbon? Mechanisms for a specific stabilisation. *Plant and Soil*, *269*(1–2), 341–356. https://doi.org/10.1007/s11104-004-0907-y

Reay, D. S., Davidson, E. A., Smith, K. A., Smith, P., Melillo, J. M., Dentener, F., & Crutzen, P. J. (2012). Global agriculture and nitrous oxide emissions. *Nature Climate Change*, *2*(6), 410–416. https://doi.org/10.1038/nclimate1458

Ren, W., Banger, K., Tao, B., Yang, J., Huang, Y., & Tian, H. (2020). Global pattern and change of cropland soil organic carbon during 1901-2010: Roles of climate, atmospheric chemistry, land use and management. *Geography and Sustainability*, *1*(1), 59–69. https://doi.org/10.1016/j.geosus.2020.03.001

Rinnofner, T., Friedel, J. K., De Kruijff, R., Pietsch, G., & Freyer, B. (2008). Effect of catch crops on N dynamics and following crops in organic farming. *Agronomy for Sustainable Development*, *28*(4), 551–558. https://doi.org/10.1051/agro:2008028

Ryle, G. J. A., Powell, C. E., & Gordon, A. J. (1979). The respiratory costs of nitrogen fixation in soyabean, cowpea, and white clover: I. Nitrogen fixation and the respiration of the nodulated root. *Journal of Experimental Botany*, *30*(1), 135–144. https://doi.org/10.1093/jxb/30.1.135

Sainju, U. M., Allen, B. L., Lenssen, A. W., & Ghimire, R. P. (2017). Root biomass, root/shoot ratio, and soil water content under perennial grasses with different nitrogen rates. *Field Crops Research*, *210*, 183–191. https://doi.org/10.1016/j.fcr.2017.05.029

Sainju, U. M., Singh, B. P., & Whitehead, W. F. (2002). Long-term effects of tillage, cover crops, and nitrogen fertilization on organic carbon and nitrogen concentrations in sandy loam soils in Georgia, USA. *Soil and Tillage Research*, *63*(3–4), 167–179. https://doi.org/10.1016/S0167-1987(01)00244-6

Salahin, N., Alam, M. K., Ahmed, S., Jahiruddin, M., Gaber, A., Alsanie, W. F., et al. (2021). Carbon and nitrogen mineralization in dark grey calcareous floodplain soil is influenced by tillage practices and residue retention. *Plants*, *10*(8), 1–17. https://doi.org/10.3390/plants10081650

Salmerón, M., Cavero, J., Isla, R., Porter, C. H., Jones, J. W., & Boote, K. J. (2014). DSSAT nitrogen cycle simulation of cover crop-maize rotations under irrigated mediterranean conditions. *Agronomy Journal*, *106*(4), 1283–1296. https://doi.org/10.2134/agronj13.0560

Singh, J., & Kumar, S. (2022). Evaluation of the DNDCv.CAN model for simulating greenhouse gas emissions under crop rotations that include winter cover crops. *Soil Research*, *60*(6), 534–546. https://doi.org/10.1071/SR21075

Smil, V. (1999). Nitrogen in crop production: An account of global flows. *Global Biogeochemical Cycles*, *13*(2), 647–662. https://doi.org/10.1029/1999GB900015

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., & Zaehle, S. (2014). Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences*, *11*(7), 2027–2054. https://doi.org/10.5194/bg-11-2027-2014

Smith, P., Calvin, K., Nkem, J., Campbell, D., Cherubini, F., Grassi, G., et al. (2020). Which practices co-deliver food security, climate change mitigation and adaptation, and combat land degradation and desertification? *Global Change Biology*, *26*(3), 1532–1575. https://doi.org/10.1111/gcb.14878

Sommer, R., & Bossio, D. (2014). Dynamics and climate change mitigation potential of soil organic carbon sequestration. *Journal of Environmental Management*, *144*, 83–87. https://doi.org/10.1016/j.jenvman.2014.05.017

Thapa, R., Mirsky, S. B., & Tully, K. L. (2018). Cover crops reduce nitrate leaching in agroecosystems:A global meta-analysis. *Journal of Environmental Quality*, *47*(6), 1400–1411. https://doi.org/10.2134/jeq2018.03.0107

Tian, H., Xu, R., Canadell, J. G., Thompson, R. L., Winiwarter, W., Suntharalingam, P., et al. (2020). A comprehensive quantification of global nitrous oxide sources and sinks. *Nature*, *586*(7828), 248–256. https://doi.org/10.1038/s41586-020-2780-0

Tian, H., Yang, J., Lu, C., Xu, R., Canadell, J. G., Jackson, R. B., et al. (2018). The global N2O model intercomparison project. *Bulletin of the American Meteorological Society*, *99*(6), 1231–1251. https://doi.org/10.1175/BAMS-D-17-0212.1

Tonitto, C., David, M. B., & Drinkwater, L. E. (2006). Replacing bare fallows with cover crops in fertilizer-intensive cropping systems: A meta-analysis of crop yield and N dynamics. *Agriculture, Ecosystems & Environment*, *112*(1), 58–72. https://doi.org/10.1016/j.agee.2005.07.003

Valkama, E., Lemola, R., Känkänen, H., & Turtola, E. (2015). Meta-analysis of the effects of undersown catch crops on nitrogen leaching loss and grain yields in the Nordic countries. *Agriculture, Ecosystems & Environment*, *203*, 93–101. https://doi.org/10.1016/j.agee.2015.01.023

Volkholz, J., & Müller, C. (2020). ISIMIP3 soil input data (v1.0). [Dataset]. ISIMIP. https://doi.org/10.48364/ISIMIP.942125

Waha, K., Van Bussel, L. G. J., Müller, C., & Bondeau, A. (2012). Climate-driven simulation of global crop sowing dates. *Global Ecology and Biogeography*, *21*(2), 247–259. https://doi.org/10.1111/j.1466-8238.2011.00678.x

Wårlind, D., Smith, B., Hickler, T., & Arneth, A. (2014). Nitrogen feedbacks increase future terrestrial ecosystem carbon uptake in an individual-based dynamic vegetation model. *Biogeosciences*, *11*(21), 6131–6146. https://doi.org/10.5194/bg-11-6131-2014

Williams, C. M., King, J. R., Ross, S. M., Olson, M. A., Hoy, C. F., & Lopetinsky, K. J. (2014). Effects of three pulse crops on subsequent barley, canola, and wheat. *Agronomy Journal*, *106*(2), 343–350. https://doi.org/10.2134/agronj2013.0274

Wittwer, R. A., Dorn, B., Jossi, W., & Van Der Heijden, M. G. A. (2017). Cover crops support ecological intensification of arable cropping systems. *Scientific Reports*, *7*(1), 41911. https://doi.org/10.1038/srep41911

Zaehle, S., & Friend, A. D. (2010). Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates. *Global Biogeochemical Cycles*, *24*(1), 1–13. https://doi.org/10.1029/2009GB003521

Zhang, B., Tian, H., Lu, C., Dangal, S., Yang, J., & Pan, S. (2017). Manure nitrogen production and application in cropland and rangeland during 1860–2014: A 5-minute gridded global data set for Earth system modeling. *Earth System Science Data*, *9*, 667–678. https://doi.org/10.5194/essd-2017-11

Zhao, J., Chen, J., Beillouin, D., Lambers, H., Yang, Y., Smith, P., et al. (2022). Global systematic review with meta-analysis reveals yield advantage of legume-based rotations and its drivers. *Nature Communications*, *13*(1), 4926. https://doi.org/10.1038/s41467-022-32464-0

Zhu, B., Yi, L., Guo, L., Chen, G., Hu, Y., Tang, H., et al. (2012). Performance of two winter cover crops and their impacts on soil properties and two subsequent rice crops in Dongting Lake Plain, Hunan, China. *Soil and Tillage Research*, *124*, 95–101. https://doi.org/10.1016/j.still.2012.05.007

Zomer, R. J., Bossio, D. A., Sommer, R., & Verchot, L. V. (2017). Global sequestration potential of increased organic carbon in cropland soils. *Scientific Reports*, *7*(1), 15554. https://doi.org/10.1038/s41598-017-15794-8

Biogeosciences

Open Access

EGU

Assessing the impacts of agricultural managements on soil carbon stocks, nitrogen loss, and crop production – a modelling study in eastern Africa

Jianyong Ma[1], **Sam S. Rabin**[1,2], **Peter Anthoni**[1], **Anita D. Bayer**[1], **Sylvia S. Nyawira**[3], **Stefan Olin**[4], **Longlong Xia**[1], and **Almut Arneth**[1,5]

[1]Institute of Meteorology and Climate Research-Atmospheric Environmental Research, Karlsruhe Institute of Technology, 82467 Garmisch-Partenkirchen, Germany
[2]Department of Environmental Sciences, Rutgers University, New Brunswick, NJ 08901, USA
[3]International Center for Tropical Agriculture (CIAT), ICIPE Duduville Campus, P.O. Box 823-00621 Nairobi, Kenya
[4]Department of Physical Geography and Ecosystems Science, Lund University, 22362 Lund, Sweden
[5]Institute of Geography and Geoecology, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany

Correspondence: Jianyong Ma (jianyong.ma@kit.edu)

Abstract. Improved agricultural management plays a vital role in protecting soils from degradation in eastern Africa. Changing practices such as reducing tillage, fertilizer use, or cover crops are expected to enhance soil organic carbon (SOC) storage, with climate change mitigation co-benefits, while increasing crop production. However, the quantification of cropland management effects on agricultural ecosystems remains inadequate in this region. Here, we explored seven management practices and their potential effects on soil carbon (C) pools, nitrogen (N) losses, and crop yields under different climate scenarios, using the dynamic vegetation model LPJ-GUESS. The model performance is evaluated against observations from two long-term maize field trials in western Kenya and reported estimates from published sources. LPJ-GUESS generally produces soil C stocks and maize productivity comparable with measurements and mostly captures the SOC decline under some management practices that is observed in the field experiments. We found that for large parts of Kenya and Ethiopia, an integrated conservation agriculture practice (no-tillage, residue and manure application, and cover crops) increases SOC levels in the long term (+11 % on average), accompanied by increased crop yields (+22 %) in comparison to the conventional management. Planting nitrogen-fixing cover crops in our simulations is also identified as a promising individual prac-

tice in eastern Africa to increase soil C storage (+4 %) and crop production (+18 %), with low environmental cost of N losses (+24 %). These management impacts are also sustained in simulations of three future climate pathways. This study highlights the possibilities of conservation agriculture when targeting long-term environmental sustainability and food security in crop ecosystems, particularly for those with poor soil conditions in tropical climates.

1 Introduction

Soils contain the largest amount of organic carbon (C) in terrestrial ecosystems, storing around 1500 Pg C (petagrams of carbon) globally (Lal, 2004). However, substantial losses of soil organic carbon (SOC) have occurred over the last decades, arising from agricultural intensification and the continuous conversion of natural soils for agricultural uses to support the food demand of a growing population (Olsson et al., 2019). The estimates of cumulative SOC loss from agricultural land vary widely, with a range of 30 to 160 Pg C across the globe for the post-1850 period (Ruddiman, 2003; Lal, 2004; Pugh et al., 2015; Smith et al., 2016; Sanderman et al., 2017). This soil carbon loss contributes to greenhouse gas (GHG) emissions in the atmosphere and thus ac-

celerates global warming. Increasing SOC storage in agricultural ecosystems through improved management practices has been repeatedly discussed as a promising option to mitigate climate change (Smith et al., 2020; Arneth et al., 2021), with co-benefits for soil fertility and crop production (Seufert et al., 2012; Knapp and van der Heijden, 2018; Shang et al., 2021).

Conservation agriculture (CA) – particularly the use of minimum soil disturbance (e.g., zero tillage), organic matter addition (e.g., crop residue retention and cover crops), and species diversification through crop rotation – is the most well-known practice to potentially enhance SOC sequestration and improve agricultural sustainability in sub-Saharan Africa (SSA; Thierfelder et al., 2013; Smith et al., 2016). Much experimental evidence has indicated that SOC stocks under CA systems are significantly higher than conventional farming practices in well-managed trials for SSA (Pittelkow et al., 2015a; Cheesman et al., 2016; Powlson et al., 2016; Sommer et al., 2018) but vary across the region due to the differences in soil properties, climate condition, and the specific management implemented in farming systems (Sanderman et al., 2017). It has been estimated that CA would increase cropland SOC by 1.2 to 2.4 Pg C over both the eastern and southern African regions if soil-conserving techniques were completely implemented over 20 years (Zomer et al., 2017). However, adoption of CA remains a challenge in the region as positive crop production effects over time can be hidden by large interannual variability in yields (Giller et al., 2009; Corbeels et al., 2014; Stevenson et al., 2014; Pittelkow et al., 2015b). Furthermore, nitrogen (N) trace gas emissions and nitrate leaching related to agricultural fertilizer also need to be investigated; these N-associated losses from agriculture have negative effects on air quality, freshwater systems, and climate from regional to global scales (Reay et al., 2012; Olin et al., 2015a; Tian et al., 2020).

Process-based ecological models with soil carbon–nitrogen (C–N) dynamics have the potential to understand and quantify the trade-offs between yields, carbon sequestration, and negative environmental effects on larger spatial scales and longer temporal perspectives due to their mathematical representation of plant growth, organic matter input, and soil decomposition (Parton et al., 1993; Li et al., 1994). These models have been widely used to explore SOC response to alternative management practices in different cropping systems (e.g., Century, Lugato et al., 2015; LPJ-GUESS, Olin et al., 2015a; RothC, Mesfin et al., 2021; LPJmL, Herzfeld et al., 2021). However, compared to temperate crop ecosystems – particularly in heavily studied North America, western Europe, and East Asia – there are limited SOC modelling studies in the tropical agroecosystems of SSA (Lemma et al., 2021; Nyawira et al., 2021). This may partially reflect the relative paucity of long-term and field-based SOC measurements in the region (Powlson et al., 2016), which limits the calibration and implementation of process-based models in assessing the impacts of land managements on SOC dynamics in the tropical SSA. In one example, Nyawira et al. (2021) examined DayCent model performance in simulating SOC and crop yield responses to improved management practices (e.g., manure and crop residue application) for maize-based cropping systems, using experimental data from two long-term field sites in western Kenya. Although model results showed a fairly good agreement with observations, the authors suggested that future model evaluation for other managements (such as cover crops) in sequestering SOC and/or reducing N losses through leaching and gaseous N emissions would be needed to support the recommendation of sustainable agricultural practices in the tropics of SSA. To date, no studies have applied process-based models on the regional scale to detect the long-term joint impacts of environmental change and alternative management practices on associated changes in crop production, C sequestration, and cropland N losses in eastern Africa, a region where agricultural soils have been experiencing strong degradation due to the combined effects of agricultural intensification and mismanagement over recent decades (Wynants et al., 2019; Mugizi and Matsumoto, 2020).

Thus, in this study, accounting for the most common and important soil-conserving techniques implemented by smallholder farms (such as conservation tillage, mineral fertilizer, and organic matter/manure incorporation), we employ a process-based dynamic vegetation model (LPJ-GUESS, Smith et al., 2014; Olin et al., 2015b) to explore and quantify the effectiveness of these alternative management practices that aim to enhance soil carbon and/or mitigate the negative effects of agriculture on the N cycle. Model results are extensively tested on experimental data from long-term (> 10 years) field trials in western Kenya and compared against country-level yield statistics, as well as region-level cropland SOC stocks from published sources. The management effects on soil C pools, crop yields, and N losses in Kenya and Ethiopia are subsequently investigated under present and future climate scenarios. The model-based and large-scale quantification of these management impacts on crop ecosystems provides a scientific understanding for identifying strategies that possibly minimize negative environmental effects while still addressing society's growing needs for food production, allowing recommendations for sustainable agricultural practices under different farming systems in the tropics of SSA.

2 Methods

2.1 Model description

LPJ-GUESS is a dynamic vegetation model with process-based representation of plant physiological and biogeochemical processes designed for regional to global applications (Smith et al., 2014). The model has been widely used to investigate vegetation and soil C–N dynamics and their inter-

actions in response to both environmental changes and management, such as changes in climate, atmospheric CO_2 concentration, N input (deposition and fertilizer rates), or irrigation. Three distinct land-use types are represented in the model: natural land, pasture, and agricultural land. Vegetation on natural land is described by the growth, disturbance, and mortality of 12 plant function types (PFTs), which differ in their bioclimatic preferences, morphological traits, and growth strategies. C_3 and C_4 grasses are modelled to represent pastures, with 50 % of aboveground biomass removed each year at harvest; the rest, together with root biomass, is assumed to return to soils as litter (Lindeskog et al., 2013). Croplands in the model are characterized by four crop functional types (CFTs, i.e., two temperate C_3 crops sown in spring and autumn, a C_4 crop representing maize, and a tropical C_3 crop representing rice), with crop-specific processes including C–N allocation, plant development stages, and explicit sowing and harvest representation at daily temporal resolution (Olin et al., 2015b). Crops in LPJ-GUESS are prescribed as either rain-fed or irrigated, with their proportions given as an external input (Lindeskog et al., 2013). Planting date is determined dynamically based on local climatology in each grid cell with five seasonality types represented (a combination of temperature- and precipitation-limited behaviours; Waha et al., 2012), and crops are harvested once every year when accumulated heat requirements are fulfilled (Lindeskog et al., 2013). At this point multi-cropping systems within a year and intercropping practices are not yet incorporated in the model. Recent relevant developments include the implementation of soil N transformation and two new legume CFTs (i.e., soybean and pulses) with symbiotic biological N fixation (BNF; Ma et al., 2022).

Soil C–N dynamics in LPJ-GUESS are simulated by a soil organic matter (SOM) scheme derived from the Century model (Parton et al., 1993), in which SOM and litter are characterized by 11 pools with prescribed C : N ratios and decay rates (Smith et al., 2014). The transfer of SOM between pools drives N mineralization or immobilization, as a result of the altered C : N ratios in the donor and receiver pool. Soil mineral N after the process of mineralization and immobilization is partially depleted by plant N uptake, which is assumed to be proportional to plant root biomass and is constrained by soil temperature, plant N status, and the mineral N pool itself (Zaehle and Friend, 2010; Wårlind et al., 2014). Leaching of mineral N is a function of the remaining nitrate concentration, percolation rate, and available soil water content. N losses through organic leaching are also included in LPJ-GUESS and associated with soil sand fraction, percolation, and the size of soil microbial SOM N pool (Smith et al., 2014; Wårlind et al., 2014). Gaseous N emission produced in the soil to the atmosphere is simulated as NH_3, NO, N_2O, and N_2, with the representation of soil N dynamic processes including ammonification, nitrification, and denitrification in the SOM pools. In this study, we combine N leaching and N gas emissions into one value to represent total N loss from

crop ecosystems. The model schematic and other calculations on cropland C–N cycles follow an earlier version of LPJ-GUESS described in Smith et al. (2014) and Olin et al. (2015b).

2.2 Alternative management practices

Agricultural management options incorporated in the model include variable sowing and harvest dates, irrigation, cover crop grass between two growing seasons, crop residue management, N fertilizer application, and tillage. The latter four practices are varied in the evaluation of management options in this study and described in detail below.

2.2.1 Cover crops

Using cover crops as "green manure" in between the main cropping seasons is an effective practice to build up or maintain soil fertility, as they can enrich soil N and soil organic carbon contents if their biomass is fully tilled into the soil. Cover crops implemented in LPJ-GUESS are modelled as C_3 and C_4 grasses grown between two agricultural growing periods of main crops, if the bare fallow duration exceeds 30 d. The cover crop leaf and root biomass are added to the surface and the soil metabolic/structural SOM pools, respectively, 15 d before the sowing date of the subsequent main crop. N-fixing herbaceous legumes such as Kenya white clover (*Trifolium johnstonii* Oliv.) and alfalfa (*Medicago sativa*) are sometimes rotated or used as intercrops between cereals to improve the soil quality in east African smallholder farming systems (Sileshi et al., 2008; Muoni et al., 2019). We thus incorporate the process of biological N fixation to C_3 grass in the model, following Liu et al. (2011), to account for the effects of herbaceous legumes on C–N cycles in crop ecosystems. The evaluation of soil C stocks, N leaching, and crop production in response to different cover crop types will be published in a forthcoming paper and therefore is not presented and discussed here. Grain legumes as cover crops are not yet implemented in LPJ-GUESS.

2.2.2 Residue retention

Leaving crop residues in field after harvest can prevent soil degradation, while also retaining water and nutrients (Smith et al., 2012). In the standard LPJ-GUESS setup, this practice is represented by removing 75 % of the aboveground biomass after harvest, thus returning the remaining 25 % of C and N mass to the soil litter pool for decomposition. In this study, we increase the residue removal fraction to 90 % for the regional simulations based on the investigated data in Ethiopia (Laekemariam et al., 2016; Lemma et al., 2021), where most smallholder farms practice mixed crop-livestock systems in which crop residues are usually removed from fields after harvest and used as fodder for livestock (Valbuena et al., 2012; Baudron et al., 2014).

2.2.3 N fertilizer and manure application

Application of N fertilizer in agricultural land is an important and widespread practice in improving crop production and enhancing SOC storage. However, if not managed appropriately, this practice can easily give rise to negative environmental impacts, like increasing soil N_2O emission (Reay et al., 2012) and/or promoting nitrate leaching to waterways (Tian et al., 2020). N fertilizer in LPJ-GUESS is applied as forms of mineral N and manure. Synthetic fertilizer application takes place at three crop development stages – sowing, halfway through the vegetative phase, and flowering – with different application rates depending on crop type (Olin et al., 2015b; Ma et al., 2022). All manure is applied to crops at the time of sowing as a single application to reflect real-world practices that account for the time required for manure N to be made available to plants. The manure application in the model is represented as N addition to metabolic and structural SOM N pools with the equal application rate. The amount of C added to soils via manure is then computed assuming a prescribed C : N ratio. The default manure C : N value of 30 (Olin et al., 2015a) was chosen to represent the C and N content from sources ranging from poultry waste (C : N of ca. 15) to straw-rich manure from livestock (C : N of 40 or more). Here, we adjust the C : N ratio of farmyard manure to 16 in all the model experiments, following the literature-based value for smallholder farming systems in eastern Africa (Gichangi et al., 2006; Nyawira et al., 2021).

2.2.4 Tillage

Different forms of tillage have been used to increase the release of nutrients from the soils for uptake by crops, but the mechanical disturbance of the soil profile increases soil erosion and heterotrophic respiration and thus enhances soil C losses to the atmosphere (Chatskikh et al., 2009; Badagliacca et al., 2018). Tillage is implemented in the model using a tillage factor, which accelerates the soil decomposition on agricultural land in the surface microbial and humus SOM C pools and the microbial and slow turnover C pools of the soil. To account for the long-term effects on heterotrophic respiration (Pugh et al., 2015; Olin et al., 2015a), the tillage factor is assumed to be a fixed value of 1.94, which is taken from Chatskikh et al. (2009) and used to modify the decay rate of the four SOM pools throughout the year.

2.3 Experimental setups

Our study is divided into three parts. In the first part we examine the model's ability to simulate the SOC and maize yield response to various managements by comparing with observed data from two long-term field sites in Kenya. Next, we update the growth parameters for sorghum in the model to better represent the agricultural production in eastern Africa because of this widely grown crop in the region. Yields for

six crop types, including the new sorghum parameterization (see Sect. 2.3.2 below), are evaluated against FAO-based statistics in Kenya and Ethiopia. In the last part, the isolated effects of each alternative management practice are first investigated for the historical period and subsequently explored under future climate scenarios by forcing the model with simulated climate over the 21st century from five general circulation models (GCMs, Eyring et al., 2016).

In order to build up cropland soil C and N pools, all simulations were initialized with a 500-year spin-up using atmospheric CO_2 from 1901 and repeated de-trended 1901–1930 climate (see Table 1 for data information). During spin-up, potential natural vegetation (PNV) was simulated for the first 470 years, and then the cropland fraction linearly increased from zero to the first historic value (1901) during the last 30 years of spin-up. Model input data are summarized in Table 1, with the different experiment setups explained in detail below.

2.3.1 Model evaluation at site scale

To evaluate the model performance, we use data from two long-term experimental sites (INM3 and CT1) managed by the International Center for Tropical Agriculture (CIAT) since 2003. The INM3 trial (0.14° N, 34.40° E) is designed to study soil fertility effects of manure and maize residue retention under conventional tillage systems, while the CT1 trial (0.13° N, 34.41° E) mainly evaluates the combined effects of conservation tillage and residue application on SOC dynamics in maize-based cropping systems (Sommer et al., 2018). A total of 16 different trials for the period 2003–2015 in a continuous maize system (two cropping periods a year) at the INM3 site were simulated: 0 and 4 t ha^{-1} (dry matter) of manure application with 2 t ha^{-1} maize residue retention or removal under four treatments of mineral N fertilizer addition (0, 30, 60, and 90 kg N ha^{-1}). Similar simulations over the same period were performed at the CT1 site, but with the difference that minimum and conventional tillage are dominant practices, and no trials receive any manure application. At present, a double-cropping system within a year has not yet been implemented in LPJ-GUESS (Olin et al., 2015a) since the second "short rainy" growing season is – from a yield perspective – not hugely relevant for most regions of eastern Africa (Wainwright et al., 2019). In this study, the second growing period under the continuous maize systems was modelled as a non-N-fixing cover crop with all the aboveground biomass removed from the field at both sites. To parameterize the N application and residue retention practices in the model, the application rate of 4 t ha^{-1} of manure dry matter was converted to 70 kg N ha^{-1} by assuming an N content of 1.75 % in farmyard manure with a fixed C : N ratio of 16 (Gichangi et al., 2006; Nyawira et al., 2021). The residue management with 2 t ha^{-1} retention was set to 50 % of maize residue left in the field, following the reported proportion described in Sommer et al. (2018) and Nyawira et al.

Table 1. Summary of simulations performed in this study. See methods section for abbreviations and further explanations.

Purpose of simulation	Code	Trial involved	Time period	Model spin-up	Land use	Climate	Manure input	Mineral N input	Number of simulations
SOC storage and maize yield evaluation against field-based trials	A1	INM3 site	1901–2015	500 years. Land use started 470 years PNV, then 30 years cropland ramp to 1901 CO$_2$ fixed in 1901	1901–2002: 100 % grassland 2003–2015: 100 % cropland	GSWP3-W5E5	70 kg N ha^{-1}	0, 30, 60, and 90 kg N ha^{-1}	16[a]
	A2	CT1 site	1901–2015	The same as A1	1901–1991 and 1995–2000: 100 % grassland 1992–1994 and 2001–2015: 100 % cropland	GSWP3-W5E5	0 kg N ha^{-1}	0, 30, 60, and 90 kg N ha^{-1}	16[a]
Regional crop yields comparison and Response to different management practices, historical	B1	F_{std} (standard simulation)	1901–2014	The same as A1, baseline simulation for B2	LUH2	CRUJRA	Zhang et al. (2017)	Ag-GRID	1
	B2	All managements[b]	2015–2100	Starting from B1 in 2014	LUH2, fixed in 2014	CRUJRA, 1995–2014 climate repeated until 2100[c]	Zhang et al. (2017), fixed in 2014	Ag-GRID, fixed in 2014	7
Response to different management practices, future	C1	F_{std} (standard simulation)	1901–2014	The same as A1, baseline simulation for C2 and C3	LUH2	Five GCMs[d]	Zhang et al. (2017)	Ag-GRID	5
	C2	All managements[b]	2015–2100	Starting from C1 in 2014	LUH2, fixed in 2014	Five GCMs × three SSPs	Zhang et al. (2017), fixed in 2014	Ag-GRID, fixed in 2014	105
	C3	All managements[b]	2015–2100	Starting from C1 in 2014	LUH2, fixed in 2014	Five GCMs, 1995–2014 climate repeated until 2100[e]	Zhang et al. (2017), fixed in 2014	Ag-GRID, fixed in 2014	35

[a] For details of 16 simulations, see Table 2. [b] All managements denote the seven practices listed in Table 3. [c] Historical (CRUJRA-based) climate with temperature de-trended. These 20 years are repeated throughout the period 2015–2100. [d] Five GCMs – GFDL-ESM4, UKESM1-0-LL, MPI-ESM1-2-HR, IPSL-CM6A-LR, and MRI-ESM2-0 – are used. GCM climate is bias corrected and statistically downscaled against observational data set GSWP3-W5E5 (Lange, 2019) at daily temporal resolution for the entire globe. [e] The same as "c", but GCM-based historical climate.

Table 2. Site- and treatment-specific data used for model evaluation at the INM3 and CT1 long-term (2003–2015) trials. The "x" in the treatment names denotes any mineral N application rate of 0, 30, 60, and 90 kg N ha^{-1}. A total of 33 % and 66 % of the mineral N fertilizer are applied at the time of sowing and halfway through the vegetative stage, respectively, following the description in Sommer et al. (2018). Abbreviations: NoMan – no manure application; NoRR – no residue retention; NoTill – no-tillage; Man – 70 kg N ha^{-1} of manure application converted from 4 t ha^{-1} dry matter; RR – 50 % of residue retention; Till – Tillage.

Site and its soil physical properties	Treatment name	Tillage	Manure (kg N ha^{-1})	Residue retention (%)	Mineral N application (kg N ha^{-1})				
					N app. timing	N0	N30	N60	N90
INM3	Nx_NoMan_NoRR	Yes	No	No	Sowing	0	10	20	30
(34.40° E, 0.14° N)	Nx_NoMan_RR	Yes	No	50	Halfway through the	0	20	40	60
Topsoil (0–20 cm):	Nx_Man_NoRR	Yes	70	No	vegetative stage				
Sand: 26 %	Nx_Man_RR	Yes	70	50					
Silt: 18 %									
Clay: 56 %									
Bulk density: 1.1 g cm^{-3}									
CT1	Nx_NoTill_NoRR	No	No	No	The same as INM3				
(34.41° E, 0.13° N)	Nx_NoTill_RR	No	No	50					
Topsoil (0–40 cm):	Nx_Till_NoRR	Yes	No	No					
Sand: 16 %	Nx_Till_RR	Yes	No	50					
Silt: 15 %									
Clay: 69 %									
Bulk density: 1.1 g cm^{-3}									

(2021). In addition, we switched off (on) the tillage option in the model to represent the minimum (conventional) tillage experiment at CT1. A summary of these trials is available in Table 2.

The gridded daily climate data set from GSWP3-W5E5 (Dirmeyer et al., 2006; Lange, 2019; Cucchi et al., 2020) at 0.5° resolution was used, and the grid cell with coordinates 34.25° E and 0.25° N is representative for the two experimental sites. We compared the required input variables from GSWP3-W5E5 with site-based weather observations, finding that the gridded climate data had a fairly good agreement with field records, although precipitation diverged between two data sets on individual days over the experimental period (Fig. S1 in the Supplement). There was not much information available on the land use in years prior to the field experiments. Therefore, to maintain soil N and C pools in equilibrium after model spin-up, we followed the simulation setups in Nyawira et al. (2021) and assumed that INM3 was under grassland systems for the period 1901–2002 with all the aboveground biomass returned to the soils (A1, Table 1), while at CT1 grassland was simulated from 1901 to 1991. After this the land use for CT1 trials was implemented according to information provided in the literature (Sommer et al., 2018): rain-fed maize farming systems from 1992–1994 (unfertilized), followed by a crop-fallow period of 1995–2000 (grassland), then 2 years with fertilized maize (18 kg N ha^{-1}) until 2002 (A2, Table 1). The modelled SOC stocks from 1901–2002 at both sites are given in Fig. S2 in the Supplement. In addition, soil physical properties in the topsoil at both sites, such as clay content (%) and bulk density (g cm^{-3}), were taken from Sommer et al. (2018) and

used as external inputs to further calculate corresponding soil hydraulic properties in LPJ-GUESS (Olin et al., 2015a).

Model performance was assessed comparing the simulated and observed maize yields and SOC stocks in response to varying management practices. For SOC comparison, the measured SOC values were scaled to 0–150 cm from the original depth (0–15 cm) to match the modelled soil depth, using the empirical depth distribution functions proposed by Jobbágy and Jackson (2000):

$$Y = 1 - \beta^d, \tag{1}$$

$$SOC_{150} = \frac{1 - \beta^{150}}{1 - \beta^{15}} \times SOC_{15}, \tag{2}$$

where Y is the cumulative proportion of the SOC pool from the surface to depth d (cm), and β is the relative rate of decrease in SOC stock with depth depending on the measured SOC content along the soil profile. The value of β is obtained from the existing literature and set as 0.971 for INM3 and 0.974 for CT1 (Nyawira et al., 2021); SOC_{15} and SOC_{150} represent the cumulative SOC stock (Mg C ha^{-1}) at 0–15 and 0–150 cm, respectively.

2.3.2 Regional crop yield evaluation

In this study we performed simulations with six CFTs – maize, pulses (representing faba bean and common bean), sorghum, wheat, rice, and soybean – which are grown widely in Kenya and Ethiopia (FAOSTAT, 2021). In a previous modelling study (Olin et al., 2015a), sorghum in LPJ-GUESS was simulated as the maize CFT. Here, we developed a CFT that better represents allocation to sorghum organs based on the

data from Penning de Vries et al. (1989) (Fig. S3 and Table S1 in the Supplement). The performance of the model for sorghum and five other crops was evaluated by comparing the simulated and reported yields at country level. For regional comparison, the crop yield statistics were collected from FAOSTAT (2021) while the simulated gridded yield (B1, Table 1) was aggregated to the national level using land-use maps (described below):

$$
\text{Yield}_{\text{country}} = \\
\frac{\sum_{i=1}^{n} [(\text{Yield}_{\text{rain}})_i \times (\text{Area}_{\text{rain}})_i + (\text{Yield}_{\text{irri}})_i \times (\text{Area}_{\text{irri}})_i]}{\sum_{i=1}^{n} [(\text{Area}_{\text{rain}})_i + (\text{Area}_{\text{irri}})_i]}, \quad (3)
$$

where $\text{Yield}_{\text{country}}$ is the aggregated yield in Kenya or Ethiopia; i is the number of grid cells in that country, varying from 1 to n; $\text{Yield}_{\text{rain}}$ and $\text{Yield}_{\text{irri}}$ denote the modelled yield under rain-fed and irrigated conditions, respectively; and $\text{Area}_{\text{rain}}$ and $\text{Area}_{\text{irri}}$ are the CFT-specific rain-fed and irrigated areas used in simulations, respectively (Fig. S4a in the Supplement).

As climate input, monthly data at 0.5° resolution from 1901–2014 were taken from observation-based CRUJRA v2.1 (Harris et al., 2020; Kobayashi et al., 2015). Annual atmospheric CO_2 concentration over the same period was from the data set provided by Meinshausen et al. (2020). Land use and land cover information was used from LUH2 (Land-Use Harmonization 2; Hurtt et al., 2020) with fractions of natural vegetation, pasture, and cropland at each grid cell, spanning from 1901 to 2014 and remapped to the same resolution of climate forcing. The fractional cover of various crop species in the year 2000 came from MIRCA (Monthly Irrigated and Rain-fed Crop Areas; Portmann et al., 2010) and was aggregated to the six CFTs modelled in this study. Generally, the total cropland cover in a grid cell could change annually over time, but the relative fractions of each CFT within that cover fraction were held constant. In addition, the soil fractions of sand, silt, and clay in the topsoil (0–30 cm) from GGCMI phase 3 (Global Gridded Crop Model Intercomparison; Volkholz and Müller, 2020) were used to parameterize soil hydraulic properties at each modelled grid cell.

Monthly atmospheric N deposition (NH_x, NO_y) for 1901–2014 was used as simulated by CCMI (NCAR Chemistry-Climate Model Initiative). The value was interpolated to 0.5° × 0.5° from the original resolution (1.9° × 2.5°) to match the resolution of the climate data (Tian et al., 2018). In terms of N fertilizer input to the cropland, CFT-specific data for mineral N fertilizer and manure over 1901–2014 came from Ag-GRID (AgMIP GRIDded Crop Modeling Initiative; Elliott et al. (2015) and Zhang et al. (2017), respectively) (Fig. S4b in the Supplement). Details on the fractions of mineral fertilizer applied to different crop development stages are provided in Table S1.

2.3.3 Ecosystem responses to management practices in eastern Africa

To detect the effects of agricultural practices on food security and environmental sustainability regionally, five alternative management practices – N-fixing cover crop ($F_{\text{CC-BNF}}$), non-N-fixing cover crop ($F_{\text{CC-NoBNF}}$), residue retention (F_{RR}), manure application (F_{Man}), and no tillage (F_{NT}) – together with an integrated management were assessed (Table 3); the latter integrated management with most individual practices included was selected to represent conservation agriculture (F_{conserv}). Simulated outputs of these six practices were compared with a conventional management prevalent in eastern Africa (F_{std}) with standard setups shown in Table 3. The practice that produced the largest SOC increase at each grid cell was chosen as the optimal soil C management (F_{opt}) for the historical and future simulations:

$$
F_{\text{opt}} = \{\text{MAX}(\text{SOC}_i - \text{SOC}_{F_{\text{std}}}), \quad i = 1 : 5\}, \quad (4)
$$

where F_{opt} is the calculated optimal (i.e., best performing) C management in a given grid cell; i represents the five management practices of $F_{\text{CC-BNF}}$, $F_{\text{CC-NoBNF}}$, F_{RR}, F_{Man}, and F_{NT}. SOC_i and $\text{SOC}_{F_{\text{std}}}$ are the modelled SOC stocks from these five practices and conventional management, respectively.

An initial experiment (B, Table 1) was performed to simulate the effects of these management practices under constant climate, CO_2, and land use in order to isolate management effects from environmental change impacts. This began with a run of the historical period (1901–2014) after model spin-up, using time-dependent gridded climate, land cover, and N inputs (deposition and fertilizer) at 0.5° resolution, combined with CO_2 concentration described in Sect. 2.3.2. The result of this run was to generate present-day cropland soil C and N pools under F_{std} over eastern Africa (B1, Table 1). Subsequent runs, one using each management practice, branched from this present-day state in 2015. In these simulations detrended climate (repeating 1995–2014) and fixed CO_2 concentration (∼ 397 ppm), together with N fertilizer and land cover data of the year 2014, were repeated for 86 years to allow soil C and N pools to reach a new equilibrium after the management shift (B2, Table 1). Our aim here was not to realistically reproduce the size of soil C and N pools in 2100 with different management practices, but rather to assess potential long-term management effects on crop ecosystems relative to the conventional practice (i.e., F_{std}). All simulated outputs in the last 10 years of the model experiments were taken for analysis.

In a second experiment (C, Table 1), simulations were driven with future monthly climate data taken from five GCMs (Eyring et al., 2016), for 1901–2100 at 0.5° spatial resolution (see Table 1 for each GCM information). The climate data were used for the entire simulation period to avoid any inconsistency between the historic and future periods. For the historical period (1901–2014), the management setup

Table 3. Simulation setups used for comparison of SOC sequestration, crop yields, and cropland N losses with different managements over eastern Africa for historical and future runs (see Sect. 2.3.3).

Simulation*	$F_{CC\text{-}BNF}$	$F_{CC\text{-}NoBNF}$	F_{RR}	F_{Man}	F_{NT}	$F_{conserv}$	F_{std}
N-fixing cover crop	Yes	No	No	No	No	Yes	No
Non-N-fixing cover crop	No	Yes	No	No	No	No	No
Residue retention	10 %	10 %	100 %	10 %	10 %	100 %	10 %
Manure application	Yes	Yes	Yes	No	Yes	Yes	Yes
Mineral N fertilizer	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tillage	Yes	Yes	Yes	Yes	No	No	Yes

* Abbreviations: CC-BNF – N-fixing cover crop; CC-NoBNF – non-N-fixing cover crop; RR – residue retention; Man – manure application; NT – no tillage; conserv – conservation agriculture; std – standard simulation, representing a conventional management prevalent in eastern Africa.

was the same as the simulation of B1 as described above, but with GCM-based climate forcings (C1, Table 1). The seven management practices listed in Table 3 started in the year 2015, with dynamic climate, CO_2 concentration, and N deposition throughout. Land cover and fertilizer use (mineral N and manure) were fixed from 2014 onwards to exclude their effects on cropland SOC sequestration (C2, Table 1). N deposition and climate data for SSP1-RCP2.6 (SSP1-26), SSP3-RCP7.0 (SSP3-70), and SSP5-RCP8.5 (SSP5-85) radiative forcing projections were selected due to the contrasting climate change and CO_2 concentration of the three scenarios (Meinshausen et al., 2020; see also Fig. S4c in the Supplement). Similar to the B2 simulation, the long-term management effects excluding environmental change impacts were also investigated using GCM-based repeated climate (C3, Table 1). The modelled SOC in the last 10 years of the C2 simulation (2091–2100) was taken to compare with the C3 output over the same period in order to explore the potential transition of the optimal soil C management (F_{opt}) caused by future climate change. Details on experimental setups are provided in Table 1.

2.4 Data analysis

The accuracy of the model in predicting SOC and yields was assessed using the coefficient of determination (adjusted R^2), relative bias (RB), absolute bias (AB), and root mean square error (RMSE):

$$RB = \frac{M_i - O_i}{O_i} \times 100\%, \quad (5)$$

$$AB = \frac{|M_i - O_i|}{O_i} \times 100\%, \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(M_i - O_i)^2}, \quad (7)$$

where M_i and O_i indicate modelled and observed values, and n is the number of observations. To evaluate the agreement of the interannual variability of modelled and reported yields in the long term, the Pearson correlation coefficient (r) was

calculated:

$$r = \frac{\sum_{i=1}^{n}(M_i - \overline{M})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{n}(M_i - \overline{M})^2 \sum_{i=1}^{n}(O_i - \overline{O})^2}}, \quad (8)$$

where $\overline{M}$ and $\overline{O}$ represent modelled and observed mean, and n is the number of years.

3 Results

3.1 Model performance at site scale

The simulated maize yields in the long rainy season (from March to August) tended to be somewhat higher than the measurements, with the mean overestimation ranging from 18 % to 21 % at the two experimental sites (Fig. 1). The averaged yields over the entire experimental period (2004–2015) between simulations and observations compared well across all the evaluated treatments, with the simulated values falling within the range of measured standard deviation (Fig. S5 in the Supplement). However, LPJ-GUESS did not capture the interannual variations in the yields well, producing a low Pearson correlation coefficient (r) and high absolute bias (AB) in all the INM3 and CT1 experiments (Table S2 in the Supplement). As expected, measured and simulated yields in combined conservation managements (e.g., manure with residue retention at INM3) were higher than the individual ones in the little-fertilized treatments, but yield discrepancies between managements became small and insignificant when maize received a high N application rate of $90\,kg\,N\,ha^{-1}$ (Table S2).

The simulated SOC at both sites showed a declining trend from 2004–2015 under all the assessed treatments, agreeing well with the observation of soil C loss over the same period; however, the model generally underestimated SOC at the beginning of experiment while overestimating soil C stocks in the last 2 sampling years (Figs. 2 and 3). A linear correlation ($p < 0.01$) between the simulated and measured SOC
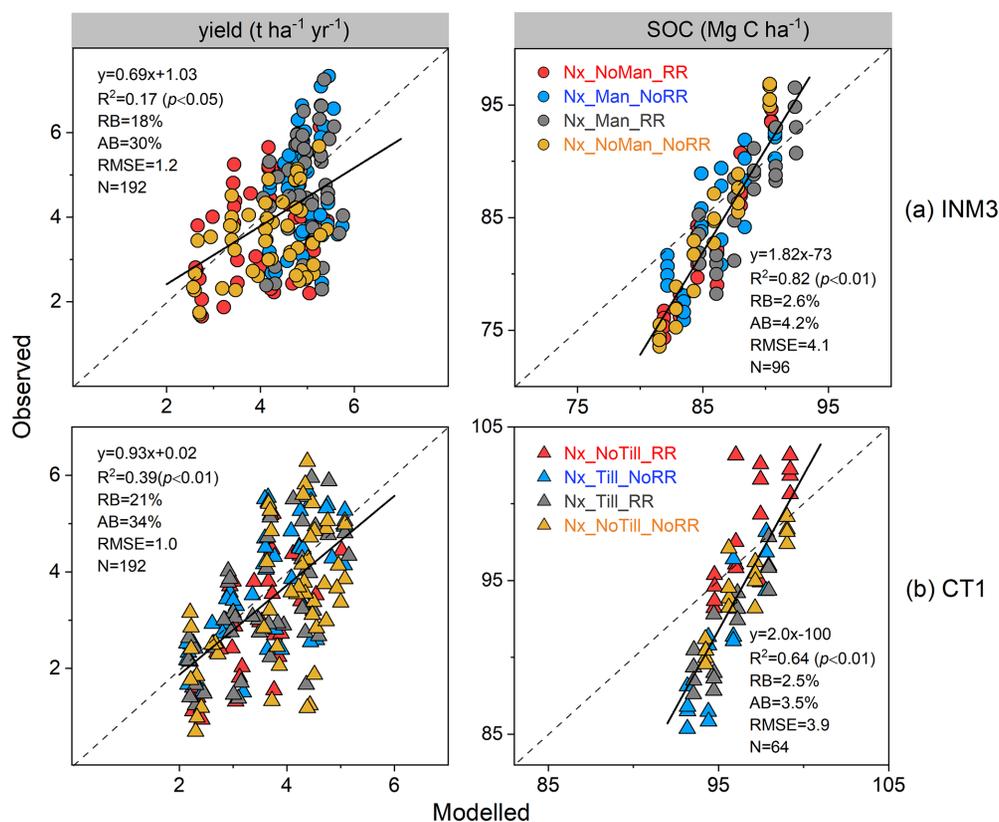
Figure 1. Comparison of modelled and observed maize yields (long rainy season, i.e., the main growing period) and SOC stocks at INM3 **(a)** and CT1 **(b)** sites for all treatments listed in Table 2. The dashed line is the 1 : 1 line, and the black bold line is the fitted linear regression; RB and AB are relative bias (Eq. 5) and absolute bias (Eq. 6), respectively, represented in percent (%); RMSE is root mean square error, with the unit of $t\,ha^{-1}\,yr^{-1}$ for yield and $Mg\,C\,ha^{-1}$ for SOC. See Table 2 for the treatment abbreviations and their explanations.

was found when all the managements were included, with the model explaining 82 % and 64 % of the variation in observed SOC at INM3 and CT1, respectively (Fig. 1). Low absolute bias of 4.2 % and RMSE value of $4.1\,Mg\,C\,ha^{-1}$ were found for the INM3 treatments, and 3.5 % and $3.9\,Mg\,C\,ha^{-1}$ for the CT1 treatments (Fig. 1). The field measurements showed that SOC stocks from the combined conservation managements were significantly higher than the conventional ones (i.e., Nx_NoMan_NoRR at INM3 and Nx_Till_NoRR at CT1). The model can generally capture this response well, but it had difficulty in predicting SOC difference between the individual managements at both sites (Figs. 1–3).

Compared to observations, LPJ-GUESS underestimated absolute SOC loss in the INM3 experiments (Table 4). Due to the extra C input to soils from manure and residue retention, the simulated combination of these two managements yielded the lowest loss of $6.7\,Mg\,C\,ha^{-1}$ at the INM3 site (N0_Man_RR), with this loss reduced by mineral N addition ($6.3\,Mg\,C\,ha^{-1}$, N90_Man_RR). By contrast, the model produced the largest C loss of $8.9\,Mg\,C\,ha^{-1}$ in the unfertilized maize control treatment with no manure and residue application (N0_NoMan_NoRR). This estimate was also similar to the simulated loss of $8.6\,Mg\,C\,ha^{-1}$ in maize

residue only (N0_NoMan_RR) and manure application only (N0_Man_NoRR). In general, the application of manure and residue retention, together with $90\,kg\,N\,ha^{-1}$ of fertilizer input, was modelled to reduce SOC loss by 29 % in comparison with the control treatment, lower than the observed reduction of 43 % (Table 4).

Similar to INM3, the observed absolute SOC loss across all the CT1 treatments was underestimated but with smaller absolute differences (Table 4). As expected, LPJ-GUESS simulated a high C loss of 4.6 and $4.5\,Mg\,C\,ha^{-1}$ in the tilled cropping systems with no maize residue retained (N0_Till_NoRR and N90_Till_NoRR). Implementing minimum tillage reduced this loss to 4.4 and $4.3\,Mg\,C\,ha^{-1}$ in treatments with residue retention included (N0_NoTill_RR and N90_NoTill_RR). Compared to the control treatment (N0_Till_NoRR), the simulated application of maize residue with $90\,kg\,ha^{-1}$ of N fertilizer reduced SOC loss by 2 % (N90_Till_RR), while adopting minimum tillage could further reduce the loss by 7 % (Table 4).
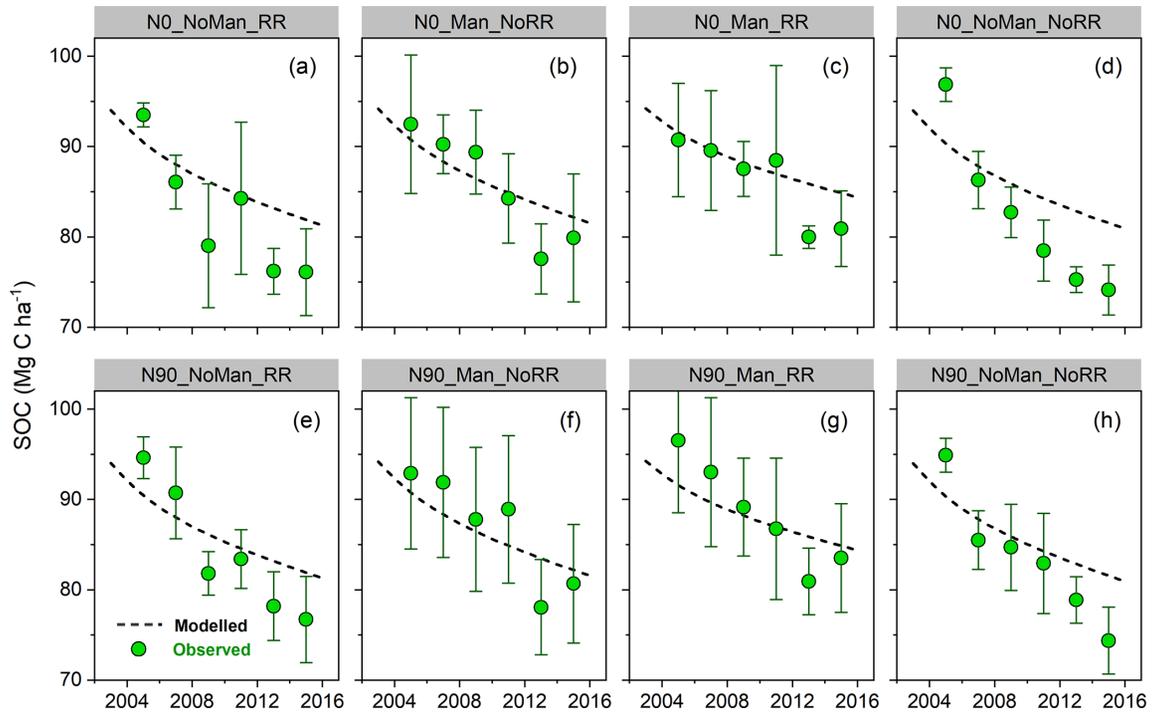
Figure 2. The modelled and observed SOC stocks (0–150 cm) for the evaluation treatments **(a–h)** at the INM3 site, with two levels of mineral N fertilizer input (N0 and N90). The dashed line is the modelled SOC, and the closed circle represents the observed value (scaled to 150 cm depth) averaged by the four replicates in the trials, with standard deviation as given in the vertical bar.
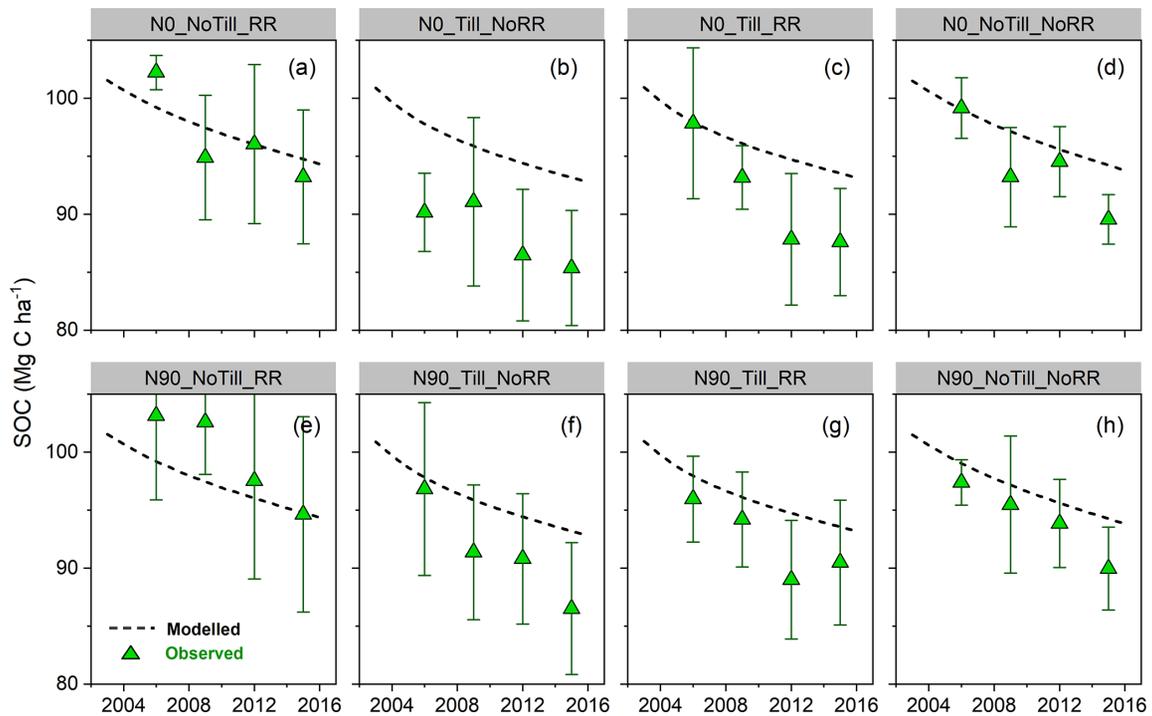
Figure 3. The same as Fig. 2, but at the CT1 site. See Table 2 for the treatment abbreviations and their explanations.

Table 4. Comparison of modelled and observed SOC stocks and absolute SOC loss for all the evaluated treatments over 2005–2015 at the INM3 and CT1 sites. The absolute SOC loss of each treatment was calculated as the difference between the first (years 2005 and 2006 for the INM3 and CT1 trials, respectively) and last sampling years (i.e., 2015). The observed SOC is represented as mean $\pm$ 1 standard deviation, deriving from the four replicates in each treatment.

	SOC ($MgCha^{-1}$)				Absolute SOC loss ($MgCha^{-1}$)	
	Year 2005 (or 2006)		Year 2015		2015 minus 2005 (or 2006)	
	Observed	Modelled	Observed	Modelled	Observed	Modelled
INM3 site						
N0_NoMan_RR	93.5 ± 1.3	90.5	76.1 ± 4.8	81.9	−17.4	−8.6
N0_Man_NoRR	92.5 ± 6.7	90.8	79.9 ± 7.1	82.2	−12.6	−8.6
N0_NoMan_NoRR	96.9 ± 1.9	90.3	74.1 ± 2.8	81.4	−22.8	−8.9
N0_Man_RR	90.7 ± 6.3	91.6	80.9 ± 4.2	84.9	−9.8	−6.7
N90_NoMan_RR	94.6 ± 2.3	90.5	76.7 ± 4.8	82.0	−17.9	−8.5
N90_Man_NoRR	92.9 ± 8.4	90.9	80.7 ± 6.6	82.4	−12.2	−8.5
N90_NoMan_NoRR	94.9 ± 1.9	90.4	74.4 ± 3.7	81.6	−20.5	−8.8
N90_Man_RR	96.5 ± 8.0	91.5	83.5 ± 6.0	85.2	−13.0	−6.3
CT1 site						
N0_NoTill_RR	102.2 ± 1.5	99.2	93.2 ± 5.8	94.8	−9.0	−4.4
N0_Till_NoRR	90.2 ± 3.4	97.8	85.3 ± 4.9	93.2	−4.9	−4.6
N0_Till_RR	97.8 ± 6.5	98.0	87.6 ± 4.6	93.5	−10.2	−4.5
N0_NoTill_NoRR	99.2 ± 2.6	99.0	89.6 ± 2.1	94.2	−9.6	−4.8
N90_NoTill_RR	103.2 ± 7.3	99.4	94.6 ± 8.4	95.1	−8.6	−4.3
N90_Till_NoRR	96.8 ± 7.5	97.8	86.5 ± 5.7	93.3	−10.3	−4.5
N90_Till_RR	95.9 ± 3.7	98.1	90.5 ± 5.4	93.6	−5.4	−4.5
N90_NoTill_NoRR	97.4 ± 1.9	99.1	89.9 ± 3.6	94.5	−7.5	−4.6

3.2 Regional yield comparison

Using the CFT-specific parameters given in Table S1, combined with the time-dependent gridded N-fertilizer data set introduced in Sect. 2.3.2, we simulated crop yields in eastern Africa under conventional management (F_{std}, Table 3) from 1901–2014. The modelled outputs from 1961–2014 and 1993–2014 were chosen to compare with annual FAO yields in Kenya and Ethiopia, respectively, due to their different time frames reported in statistics.

Modelled maize yields in the two countries showed a good agreement with observations, with a low relative bias (RB) of −6 % and RMSE value of $0.22\,tha^{-1}\,yr^{-1}$ in Kenya and −21 % and $0.54\,tha^{-1}\,yr^{-1}$ in Ethiopia (Fig. 4a and b). LPJ-GUESS tended to overestimate the reported yields in pulses and sorghum, with the country-level overestimation spanning from 48 %–257 % and 72 %–203 %, respectively. With the exception of sorghum in Kenya, the correlation between the simulated and reported yields was positively significant in most crop types, with a Pearson correlation coefficient (r) from 0.55–0.90 ($p < 0.001$, Fig. 4a and b), indicating that the model was able to capture the interannual variability in yields despite some deviations from observations for individual years. Also, the total modelled maize production on

a regional scale increased from 5.4 million tonnes in 1993 to 9.8 million tonnes in 2014, in line with the reported values of 3.6–11.2 million tonnes per year over the same period (Fig. 4c). Including all six agricultural crops in LPJ-GUESS gave a mean total production of 19.7 million tonnes per year from 1993–2014, ~ 45 % higher than the FAO statistics of 13.6 million tonnes per year, mainly due to the large overestimation in pulse production regionally (6.6 and 2.2 million tonnes per year for modelled and reported yields, respectively).

3.3 Ecosystem responses to management practices in eastern Africa

3.3.1 Historical runs

With six crop types included, all the explored management options that address aspects of sustainable land management resulted in a net increase in simulated cropland SOC in Kenya and Ethiopia compared to the conventional management (Fig. 5a). As expected, our simulation of the integrated conservation agriculture practice generated nearly the largest increase in soil C sequestration of ~ 11 %, followed by cover crops (both N-fixing and non-N-fixing), residue
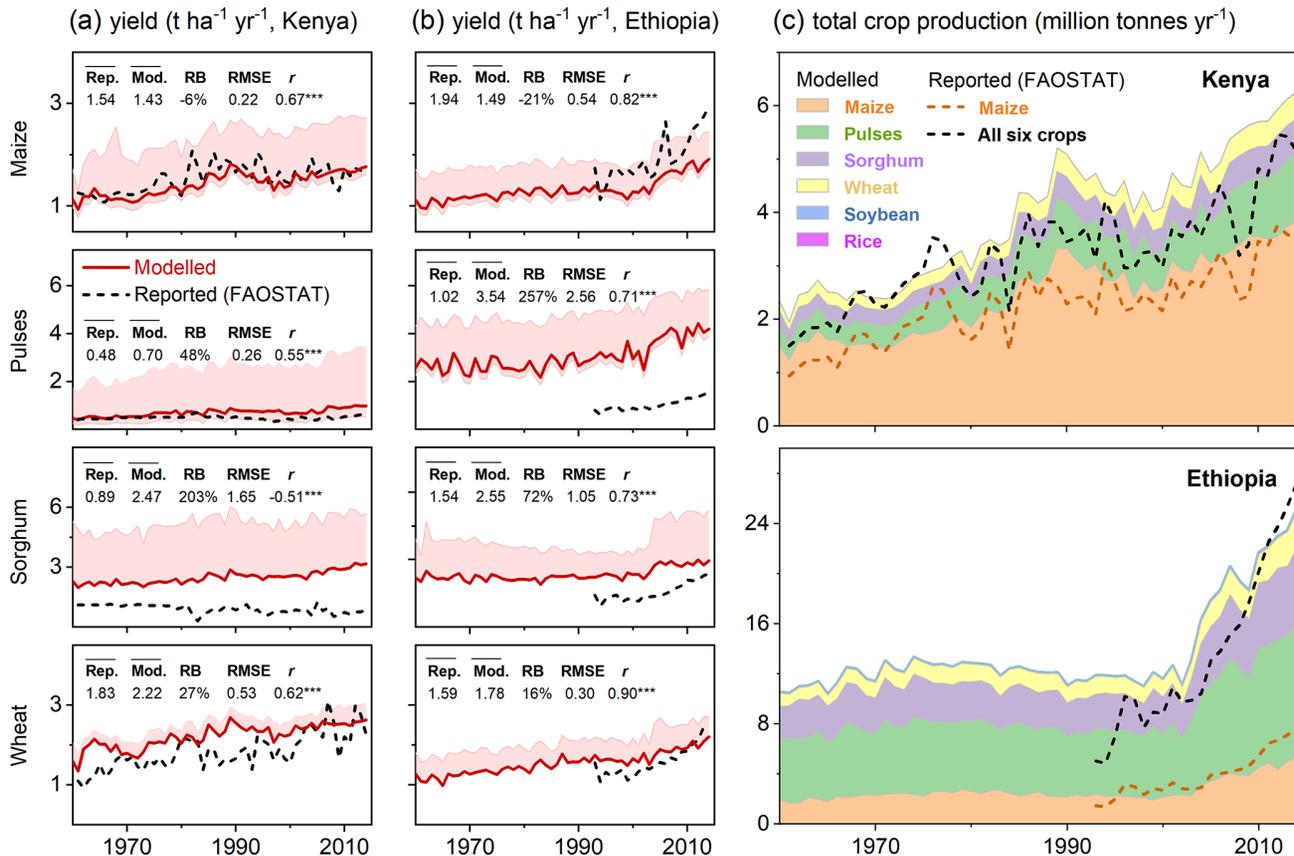
Figure 4. Comparison of modelled and FAO-reported annual crop yields at the country level from 1961–2014 in Kenya (**a**), in Ethiopia (**b**), and for total crop production (**c**). The upper and lower bounds of shade areas in (**a**) and (**b**) represent the simulated yields in irrigated and rain-fed conditions, respectively, with their area-weighted aggregated results as given in red solid lines. $\overline{\text{Rep.}}$ and $\overline{\text{Mod.}}$ indicate the reported and modelled yields averaged over FAO-based periods (1961–2014 for Kenya and 1993–2014 for Ethiopia), respectively; RB is relative bias, represented in percent (%); RMSE is root mean square error, with the same unit as yield ($\text{t}\,\text{ha}^{-1}\,\text{yr}^{-1}$); r is Pearson correlation coefficient, where *** denotes the correlation being statistically significant at the $p = 0.001$ level.

retention, and manure application, with the lowest increase of $\sim 2\,\%$ found in no-tillage management practice. Most of these investigated practices also achieved the extra benefit of increased yields – despite being accompanied by larger N losses in our simulations – with the exception of cover crops in some regions. Compared to a non-N-fixing cover crop, the implementation of a N-fixing cover crop was modelled to broadly produce higher N loss over eastern Africa. However, this practice was accompanied by an increase in simulated yields of $\sim 18\,\%$, as a result of additional N input through symbiotic N fixation in herbaceous legumes, which facilitates a N-rich soil environment to subsequent crops for better growth and productivity. Leaving all the crop residues in the field and applying manure as fertilizer were the two treatments that increased the modelled yields for most croplands but with the large environmental "cost" of an increase in N loss. The increase in both yield and N loss from residue retention likely reflects that N becomes available for crop uptake over a longer period, and nothing grows between the growing periods, which can increase the N leaching from soil. In

addition, no tillage, as an important component in conservation agriculture in the tropics of Africa, was simulated to potentially reduce the N loss from cropland with slight yield benefits depending on region (Fig. 5a) and cropping system (Fig. S6 in the Supplement).

The impacts of individual (and combined) management techniques varied widely between different parts of Kenya and Ethiopia, depending on climate and soil condition, as well as crop types (Fig. S6). In general, N-fixing cover crop was identified as a promising option for potentially sequestering SOC, with 43 % of cropland grid cells having this practice as the optimal soil C management (F_{opt}), followed by manure, residue retention, and the conventional management practice (F_{std}, Fig. 5b). However, this spatial pattern showed a distinct difference between crop types. For instance, incorporating crop residue into soil was simulated to dominate soil C responses in maize and sorghum systems, but it only slightly contributed to SOC enhancement in wheat and pulse cropping systems (Fig. S7 in the Supplement), likely reflecting their differences in biomass production, phenological re-
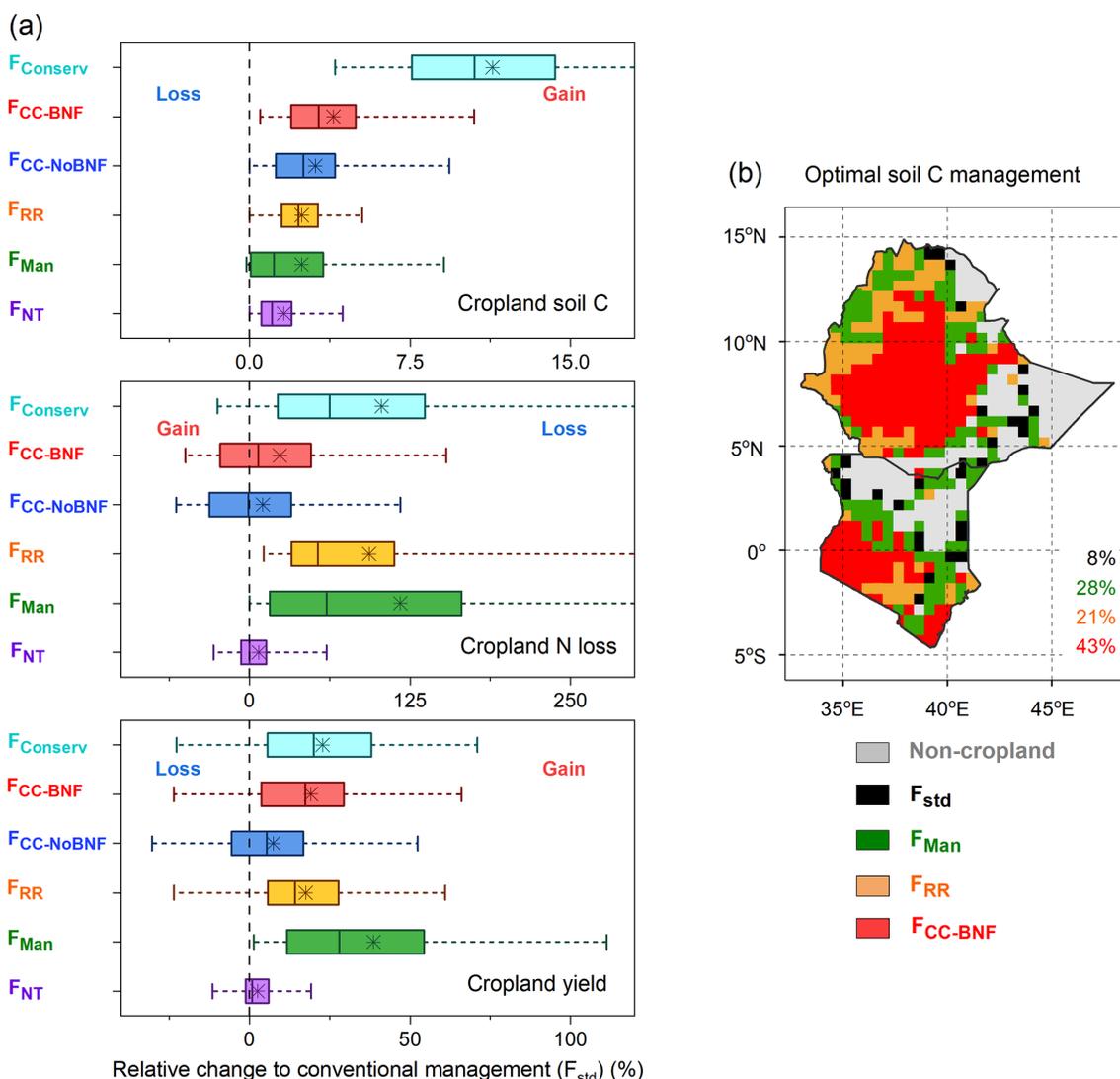
Figure 5. The modelled relative response (%) of cropland SOC, N loss, and yield to alternative management techniques (see Table 3 for abbreviations) compared to the conventional management prevalent in eastern Africa **(a)** and the optimal soil C sequestration practice (F_{opt}, Eq. 4) simulated by LPJ-GUESS in Kenya and Ethiopia **(b)**. Box plots in **(a)** denote the 5th and 95th percentiles with whiskers, median and interquartile range with box lines, and mean with asterisks across all cropland grid cells (428). The numbers in **(b)** represent the grid cell proportion of each optimal management from the total grid area. The conventional management (F_{std}, black in **b**) was chosen when no other alternative managements yielded a net increase in SOC.

sponses to climate change, and N-fertilizer application rates (Fig. S4b).

The simulated cropland soil C stock (0–150 cm) from various managements ranged from 932–1038 Tg C (teragrams) in Kenya and 2569–2895 Tg C in Ethiopia, which, as expected, was larger than the published sources for the depth layer 0–30 cm (Zomer et al., 2017). However, these modelled soil C stocks compared reasonably with the scaled-up published estimates based on the depth distribution functions (Eqs. 1 and 2), with 727 and 2227 Tg carbon estimated for the depth of 0–150 cm in Kenya and Ethiopia, respectively (Table 5). The simulated N loss of 61 Gg N yr^{-1} (gigagrams) under

the conventional management (F_{std}) in Kenya is lower than the statistic-based estimate of 111 Gg N yr^{-1} (Zhang et al., 2021), but returning all the residues to the soils (F_{RR}), the model gave a N loss of 134 Gg N yr^{-1} (Table 5), comparable to the findings in Zhang et al. (2021). Additionally, the total simulated maize production of 8.7–14.3 million tonnes per year on a regional scale was close to the FAO-reported yield of 11.2 million tonnes per year. With all agricultural crops included, an overall overestimation of 7 %–47 % was found (Table 5), primarily reflecting the overestimated production in pulses and sorghum described in Sect. 3.2 (Fig. 4).

Table 5. Modelled total cropland soil C stocks (0–150 cm), N loss, and total crop production with different management options in Kenya and Ethiopia, compared to literature-based estimates. See Table 3 for abbreviations.

Management	Soil C stock, total (TgC)		N loss (GgN yr^{-1})		Crop production (million tonnes per year)			
	Kenya	Ethiopia	Kenya	Ethiopia	Kenya		Ethiopia	
					Maize	All crops[a]	Maize	All crops[a]
F_{std}	939	2592	61	157	3.9	7.3	7.3	21.7
F_{NT}	948	2623	64	164	3.8	7.6	7.1	23.4
F_{Man}	932	2569	45	79	3.2	6.7	5.5	19.9
F_{RR}	957	2653	134	359	4.2	8.4	8.2	26.7
$F_{CC-NoBNF}$	969	2696	64	190	3.9	7.7	7.7	23.1
F_{CC-BNF}	979	2710	75	204	4.9	8.9	8.5	24.7
F_{opt}	993	2786	81	229	4.6	9.0	8.3	25.8
$F_{conserv}$	1038	2895	127	375	5.1	9.4	9.2	27.0
Other studies	414[b] 727[c]	1268[b] 2227[c]	111 (76–297)[d]	–	3.5[e]	5.2[e]	7.7[e]	19.6[e]

[a] Summed yield of six crop types: maize, pulses, sorghum, wheat, rice, and soybean. [b] Zomer et al. (2017). [c] Zomer et al. (2017); soil C stocks were scaled up to 0–150 cm from the original depth of 0–30 cm using the depth distribution functions (see Eqs. 1 and 2). [d] Zhang et al. (2021); the mean estimate over 2006–2015 was chosen, with a range given in parentheses. [e] FAOSTAT (2021); the reported total production in the year 2014 was used for comparison, since the simulated cropland area was fixed from 2014 onwards ($\sim 6.2 \times 10^6$ and 17.4×10^6 ha for Kenya and Ethiopia, respectively); see B2 in Table 1.
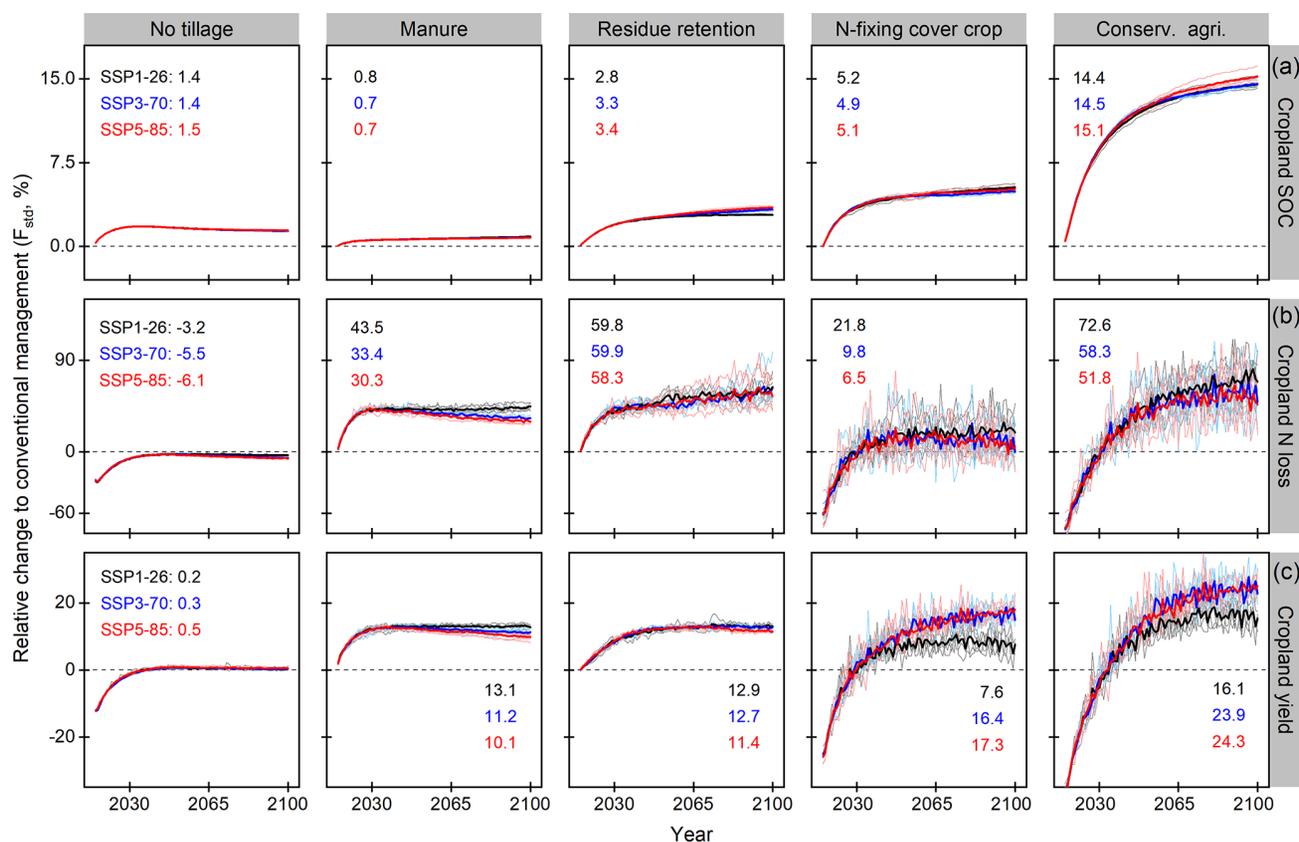
Figure 6. The simulated response (%) of cropland SOC (**a**), N loss (**b**), and yield (**c**) to alternative management techniques, relative to the conventional management (F_{std}). The dark black, blue, and red lines denote the mean of simulations using five GCMs (see Table 1) for SSP1-26, 3-70, and 5-85 scenarios, respectively. Lines in lighter colours represent the simulation driven by individual GCMs. Numbers in plots indicate the averaged results between 2091 and 2100.
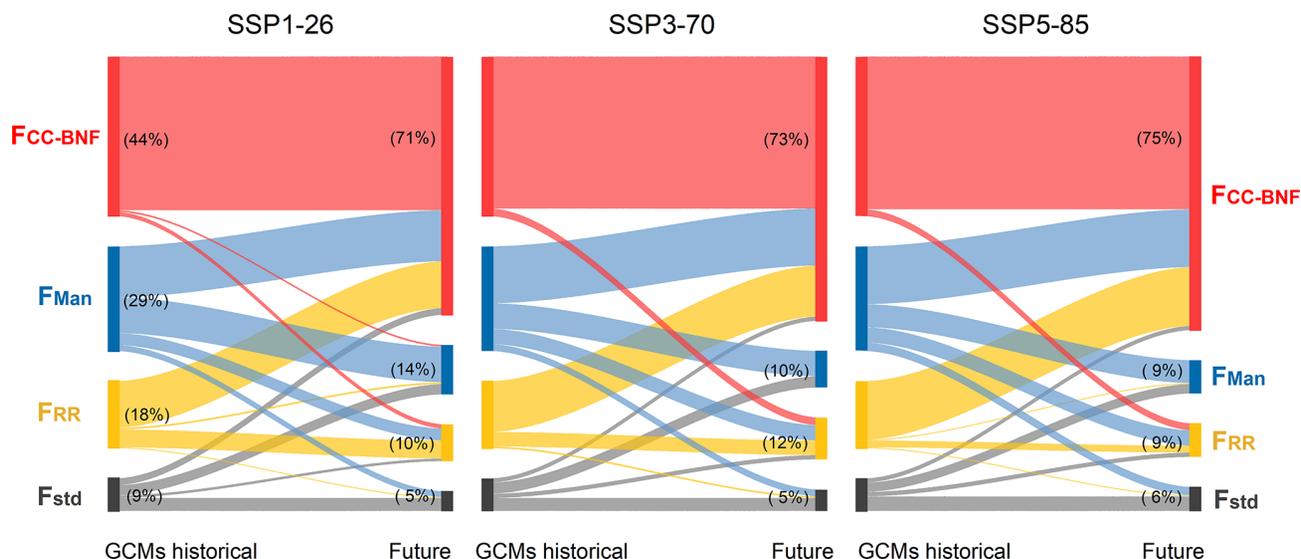
Figure 7. The relative (%) number of cropland grid cells regarding their optimal soil C sequestration practice (Eq. 4) for the historical period (C3, Table 1) and three future SSP scenarios from 2091–2100 (C2, Table 1). The numbers in the figure represent the mean of simulations using five GCMs, and transitions are indicated by the coloured bands. See Table 3 for management abbreviations.

3.3.2 Future projection

Compared to the standard model setup (F_{std}), all management practices were simulated to enhance the C storage in agricultural soils at the end of this century (i.e., 2091–2100) but with insignificant differences between three future climate change and CO_2 scenarios (Fig. 6a). Although no tillage had nearly no impact on crop production, it was accompanied by the environmental benefit of N loss reduction (Fig. 6b). A clear yield difference between the three SSP scenarios was consistently seen in experiments with N-fixing cover crop and conservation agriculture practices, with production increases being higher for SSP5-85 than for the SSP1-26 climate pathway (Fig. 6c). This likely reflects the stronger CO_2 fertilization effect on the growth of herbaceous legumes under SSP5-85. Overall, the future projection showed that N-fixing cover crop represented a near win–win situation in the sense of SOC enhancement and yield increase in eastern Africa, also with lower N loss compared to manure and residue application practices.

The simulated cropland soil C stocks (0–150 cm) under future conditions varied widely between the assessed management options, with the integrated conservation agriculture being the only practice that showed positive soil C sequestration over the simulation period (2015–2100). Adoption of N-fixing cover crops contributed to increasing SOC stocks in the first 2 decades of the model experiments, after which stable SOC for SSP1-26 and slight C loss for SSP3-70 and 5-85 scenarios were simulated (Fig. S8 in the Supplement). Other practices, such as the conventional management and no tillage, generally exhibited an obvious declining trend in total C storage between 2015 and 2100. Also, there were substan-

tial changes in the optimal C sequestration practice for the future scenarios, with $\sim 30\%$ of cropland areas in Kenya and Ethiopia (Table S3 in the Supplement) showing the potential transitions in the last 10 years of this century in comparison to the present-day climate (GCM-based historical simulation; see C3, Table 1). Most of these shifts were simulated to come from the other management type options for N-fixing cover crop, such as manure application and residue retention (Fig. 7).

4 Discussion

4.1 Uncertainties on model evaluation at site scale

LPJ-GUESS simulates the average maize yields among treatments over the experimental period well (2004–2015; see A1 and A2 in Table 1), but the measured interannual variability of the yields for the different management treatments was not well predicted. This issue is not unique to our study and has also been found at the same sites using the DayCent model (Nyawira et al., 2021). The poor performance in modelling yield variability is likely due to the precipitation discrepancy between the gridded climate input data (i.e., GSWP3-W5E5) and field-based weather records (Fig. S1), resulting in the effects of extreme weather events (e.g., drought, rainstorms, or flooding) being difficult to account for. Also, these impacts of extremes on physiological processes such as flowering or grain filling are not well represented in crop models so far, including LPJ-GUESS, but known to cause yield losses (Olin et al., 2015a; Nyawira et al., 2021).

Since LPJ-GUESS at this point does not simulate multi-cropping within a year, absence of maize residue and manure

application events in the second cropping season (i.e., the short rainy season from September until January in western Kenya) may contribute to the underestimation of the measured SOC in the treatments with these two practices included (Table 4). In addition, compared to the fixed amount of maize residue retained in the field trials ($2\,t\,ha^{-1}$), using 50 % of residue retention in the model setup will introduce some variation in terms of C inputs to soils because of the varying biomass of simulated maize residue between years (~ 1.3–2.2 and 1.4–2.4 $t\,ha^{-1}$ of residue returned to the fields in CT1 and INM3 simulations, respectively; not shown). This may partially explain the differences in the rates of SOC loss between the observed and simulated values at both sites.

All investigated management practices led to decline in SOC stocks in the field trials (Figs. 2 and 3); the overall trends were also reproduced by the model. Nonetheless, the soil C loss rates from 2004–2015 were unexpected, since addition of farmyard manure and residues can enhance SOC storage via additional C inputs to soils while conservation tillage slows down decomposition in the SOM pools. Both INM3 and CT1 sites in this study were under natural grassland before the trials start (see A1 and A2, Table 1); hence SOC losses in observations and simulations reflected (a) grassland soils tending to store more carbon than cropland (Fig. S2), and (b) a new SOC equilibrium may not have been reached in the maize cropping systems after 10+ years of cultivation (Lal, 2008). A similar finding was reported by Moebius-Clune et al. (2011), who showed declining SOC in western Kenya even after more than 50 years of conversion from primary forest to maize. Furthermore, fast turnover of the SOM in the humid tropics could be another factor affecting the SOC trends because of the prevailing warm and moist climate (i.e., western Kenya in this study). The turnover-driven C losses at the sites may exceed the gains from the C addition from manure and residue application (Kihara et al., 2020; Nyawira et al., 2021).

Agreement between the observed and simulated SOC declines was reasonable for all the considered treatments, although LPJ-GUESS generally underestimated the rates of SOC loss at the two experimental sites (Table 4). Previous studies have shown that high termite activity in western Kenya can strongly promote litter decomposition rates in the non-tilled maize cropping system (Ayuke et al., 2011; Kihara et al., 2015). We do not know whether this particular process played an important role in the field trials, but it is not included in the representation of SOM decay in the model (see Sect. 2.1). In principle, decomposition by soil animals could be addressed by adjusting the decomposition parameters in the structural and metabolic litter pools (Nyawira et al., 2021), but adopting such an approach is currently prevented by the lack of information for evaluation.

In order to compare the modelled SOC stocks with observations, as described in Sect. 2.3.1, we scaled up the measured SOC in the upper soil (0–15 cm) to the modelled depth of 150 cm using a simple extrapolation function. However, the extrapolated SOC values are most likely different from observations for the depth of 0–150 cm because of the varying management effects on SOC changes with depth. For instance, a recent analysis indicated that an intermediate and high intensity of tillage can significantly reduce SOC storage in agricultural soils, but large variations existed between soil layers (Haddaway et al., 2017). Scaling the SOC values with depth in the analysis cannot reflect this variability and introduces uncertainties on soil carbon estimates in our evaluation.

4.2   Regional yield comparison

Our simulated maize production at the country level agreed well with FAO statistics in Kenya and Ethiopia, but a general yield overestimation was found for most other crop types (Fig. 4). One factor contributing to the overestimation is that LPJ-GUESS applies a harvest efficiency of 90 % to adjust the modelled crop yields on large spatial scales (Lindeskog et al., 2013). This value has been chosen to account for the crop post-harvest losses arising from mechanical and/or manual damage during harvest operation or poor handling and/or storage conditions (Sheahan and Barrett, 2017; Stathers et al., 2020). The FAO (2011) reports that the quantity loss for cereals varies widely between regions due to the differences in management technology, ranging from 5 %– 7 % in Europe and North America to 18 % in sub-Saharan Africa (SSA). If the reported losses for SSA also apply to Kenya and Ethiopia, the value of 90 % implemented in the model would lead to a yield overestimation by 10 % regionally.

A strong overestimation in pulse production was seen for both countries (Fig. 4). This can likely be explained by the high legume N fixation capacity modelled by LPJ-GUESS in warm and moist climates (Ma et al., 2022). A high rate of BNF may reduce the N constraints on leaf photosynthesis and subsequently strengthen the flow of carbon assimilation to storage organs, resulting in high production in N-fixing crops. Yet, similar to pulses, our simulated sorghum yields at the country level were also significantly greater than FAO records (Fig. 4). This suggest that other factors are at play as well. For example, insect pests, particularly shoot flies and stalk borers, have been identified as the major constraint to sorghum production in SSA, leading to an estimated yield reduction of 11 %–49 % in western Africa and 15 %–88 % in eastern Africa (Okosun et al., 2021). LPJ-GUESS does not yet take pests into account, which could contribute to the large overestimation of sorghum production in our studied region. Additionally, a good representation of photosynthate allocation to various plant organs is important when modelling crop yields (Bondeau et al., 2007). In this study we updated the daily assimilate partitioning scheme of sorghum based on the existing literature (Fig. S3), but this process has not yet been parameterized and calibrated against observations from field experiments. Whether or not this is related

to the large-scale yield overestimation needs to be further investigated in future work.

4.3 Ecosystem responses to management practices in eastern Africa

4.3.1 Soil carbon stocks

Published estimates of management improvement effects on the potential SOC increase on Kenya and Ethiopia cropland vary between 15.9–32.7 Tg yr^{-1}, assuming that the improved managements are continuously practised over 20 years (Zomer et al., 2017). Across the eastern African study region, LPJ-GUESS predicted a SOC increase of 2.9 Tg yr^{-1} for the optimal C management (F_{opt}) and 4.7 Tg yr^{-1} for the integrated conservation agriculture practice ($F_{conserv}$) compared to the conventional management (F_{std}, Table 5). The difference between the estimates in Zomer et al. (2017) and our study may well be caused by our longer simulation period. When a change in management causes soil C stock to increase, it moves towards a new equilibrium value over a period of years or decades depending on climate and soil type (Johnston et al., 2009; Sommer and Bossio, 2014). In the early years after the change in management, the annual rate of increase is largest, and it then gradually declines when the new SOC equilibrium value is approached (Poeplau and Don, 2015; Powlson et al., 2016). The 86 years of simulations in our study are roughly 4 times longer than the 20 years studied in Zomer et al. (2017), and hence lower simulated annual rates of SOC increases are expected. If we consider the rates of SOC change over the first 20 years of simulation, the modelled soil C increase of 13.6 Tg yr^{-1} from $F_{conserv}$ practice (not shown) is close to the lower end of the range found in Zomer et al. (2017).

Most CA techniques adopted in sub-Saharan Africa (SSA) are the combined treatments of minimum tillage and residue retention (Thierfelder et al., 2013; Cheesman et al., 2016). At the regional level, our modelled small SOC increase of 2 % in no tillage (F_{NT}) and 3 % in residue retention (F_{RR}) agree with a recent meta-analysis of Githongo et al. (2021), in which converting from a conventional tillage to a no-till system in SSA on average showed only slight SOC increase in a maize cropping system. This reported insignificant impact contrasts with an earlier synthesis conducted by Powlson et al. (2016), who reported that the combination of minimum tillage and residue retention in SSA would result in a net SOC increase of 0.45 t C ha^{-1} yr^{-1} after 3 to 9 years of implementation, ~ 24 % higher than the control management (i.e., tillage and residue removal). In our model experiments, only the integrated conservation agriculture practice ($F_{conserv}$) results in a fairly large SOC increase of 11 % (varying from 4 %–22 %, Fig. 5a), more comparable but still below the findings in Powlson et al. (2016). The reason for the disagreement between the regional simulation and field-based experiments is difficult to assess because of the dif-

ference in the studied geographical scales, land use history, sampled soil depth, and implemented duration of practices. Nevertheless, they point robustly to the potential of affecting soil C storage positively through management, even though the magnitude remains unresolved.

Reflecting poor soil condition and limited manure availability, cropland SOC stocks are generally proportional to the applied amount of manure in SSA (Gross and Glaser, 2021). In our study, the regional mean application rate of manure from the year 2014 is modelled to vary from 12–32 kg N ha^{-1} yr^{-1} among crop types (Fig. S4b), resulting in an overall SOC increase of 3 %, with a range of 0.2 %–9.1 % depending on grid cell (F_{Man}, Fig. 5a). This simulated increase is comparable with results from the field experiments. For instance, in a 4-year trial, Alemu and Bayu (2005) found that the SOC in Ethiopian sorghum fields with 21 and 42 kg N ha^{-1} yr^{-1} of manure inputs was 7.8 % and 9.4 % higher than the control treatment, respectively. A similar SOC increase of 8 %–11 % was also reported under maize systems with residue removal in western Kenya, but with 140 kg N ha^{-1} yr^{-1} organic fertilizer applied for 12 years (Sommer et al., 2018). In our study we implemented a uniform C : N ratio of manure for all the simulated years and grid cells based on literature values (see Sect. 2.2.3). These set values cannot reflect the known considerable variation in C : N between manure types and locations in eastern Africa that arise from different plant species consumed by the farm animals and from different farm animal species (Zhu et al., 2020). Absence of spatial variations in C : N ratio could bias the amount of C added via manure application events and thus increase the uncertainty in the model predictions on SOC stocks.

Our modelled regional-scale results are consistent with a recent meta-analysis finding that N-fixing legume cover crops contribute more to increasing SOC storage than do non-legume plants (Abdalla et al., 2019). However, it should be noted that in LPJ-GUESS we assumed that cover crops in eastern Africa are rotated with the main crops and thus solely grown during the short rainy season. This assumption is likely to result in cover crop biomass input to the soil pools being too high as we may overestimate the length of the fallow period for cover crop growth (Porwollik et al., 2022). Such an overestimation would then also be reflected in high SOC estimates. At present more than 90 % of total annual crop yields in Ethiopia are achieved in the long rainy season (Central Statistical Agency, 2016); nevertheless, most farmers are reluctant to implement a "main crop (long rainy season) + cover crop (short rainy season)" rotation system since this practice still requires sacrificing one (short) season of maize production. Our model experiments support earlier findings that planting leguminous cover crops during the short rainy season is expected to sustainably achieve SOC and may lead to yield increases in the tropics of SSA (Rao and Mathuva, 2000; Carsky et al., 2001), although yield benefits from N-fixing cover crops at some smallholder farms

may not fully compensate for the production loss of the short rainy season (Carsky et al., 2001).

The view that adopting CA techniques can increase SOC storage in agricultural soils for SSA is based on analyses of differences between management practices, but without a time perspective (Martinsen et al., 2019; Kihara et al., 2020). The future projections done here point out that the observed SOC decline in 12-year trials in western Kenyan (Figs. 2 and 3) would continue to be found in other parts of eastern Africa under most assessed management practices with the exception of $F_{conserv}$ (Fig. S8), in line with the finding of a recent modelling study in Kenya (Nyawira et al., 2021). The 4p1000 initiative (https://www.4p1000.org/, last access: 14 November 2021) launched at COP 21 in Paris sets a target of $3.4\,\mathrm{Pg\,C\,yr^{-1}}$ SOC sequestration in agricultural soils (0–40 cm) worldwide to contribute to mitigating global climate change (Corbeels et al., 2019). Our modelling results indicate that croplands situated in eastern Africa can achieve this target only if a combination of management practices would be adopted and sustained. But even though altered management practices may not always support a positive soil C sequestration at regional scale (especially under climate change), they nonetheless are here projected to lower SOC losses and have co-benefits for crop production (Kihara et al., 2020; see also Figs. 5 and 6).

4.3.2 Cropland N loss and yields

Compared to North America, western Europe, and East Asia, annual total N loss from agricultural soils in SSA is rather small, mainly due to the low N fertilizer use across the region (Liu et al., 2010; Bouwman et al., 2013). Our model-simulated N loss of $45–134\,\mathrm{Gg\,N\,yr^{-1}}$ in Kenya is approximately half of statistics-based estimates of $76–297\,\mathrm{Gg\,N\,yr^{-1}}$, using a nitrogen-budget method (Zhang et al., 2021). Likewise, our standard model setup (F_{std}) simulated regional mean N loss of $9.2\,\mathrm{kg\,N\,ha^{-1}\,yr^{-1}}$ (Table 5), again about half of the estimated $16.7–18.2\,\mathrm{kg\,N\,ha^{-1}\,yr^{-1}}$ in eastern Africa reported by Kaltenegger et al. (2021). One possible reason for these discrepancies may be missing processes in LPJ-GUESS, such as N loss via soil erosion or surface runoff. Also, the nitrogen-budget method adopted in Zhang et al. (2021) and Kaltenegger et al. (2021) assumed that all the crop residues were retained in the soils after harvest, contrasting with the setups in our F_{std} simulation (only 10 % residue retention, Table 3). Removing most residues from cropland in the model experiment is expected to produce low N loss because of less N inputs to the soils compared to 100 % of residue retention. With all the residues left in the fields (F_{RR}, Table 3), the model regionally showed N loss of $20.8\,\mathrm{Gg\,N\,yr^{-1}}$ (Table 5), comparable to the findings in Kaltenegger et al. (2021).

For untilled maize and pulse cropping systems, we found a negative correlation between simulated crop production and N loss (Fig. S6), implying that yield increases derived from

N loss reduction are possible in eastern Africa by no-till management in the medium to long term. However, at least in our simulations, these production benefits may not be attained in the first 2 decades of simulations (Fig. 6c) because of the time that soil C and N pools need to adjust to the change in management. A similar finding also emerged from a meta-analysis of Pittelkow et al. (2015b), which pointed to increased yields, globally, in cereal and legume cropping systems only after more than 10 years of conversion from a conventional tillage to a no-till system.

An increase of 89 % in N losses was simulated across the study region in response to leaving crop residues in the field compared to the model standard setup (Fig.5a). A global analysis found that leaving residues behind might increase gaseous N emission by 8 %–37 % (Xia et al., 2018), but the same study also estimated straw return to reduce hydrological N losses by 10 %–26 %. At present the crop residue implementation and soil representation in LPJ-GUESS do not account for soil hydraulic properties in response to residue application. This missing process is likely to result in an overestimation of hydrological N losses since straw return generally reduces N leaching through enhancing soil water retention in reality (Blanco-Canqui et al., 2007). In addition, crop residues after harvest in eastern Africa are expected to rapidly decompose in response to the warm and moist climate, continuously releasing reactive N for plant uptake in the subsequent cropping period (Kihara et al., 2015). As discussed earlier, only a single growing season within a year was represented in LPJ-GUESS, and without cover crops, the relatively long bare fallow period under the simulated residue retention systems would amplify N losses as the mineralized N is not used to support plant growth during the short rainy season. Nevertheless, modelled crop production induced by retained residues still increased by 18 % on regional average (Fig. 5a) because of the enhanced size of the mineral N pool. This result is in line with two previous studies that showed a mean increase of 19 %–35 % over a 3-year period in Ethiopian wheat systems with 66 % residue retention (Adimassu et al., 2019) and another reporting $-1\,\%$–39 % of maize yield changes in response to residue retention management in semi-arid Kenya (Kihara et al., 2011).

In some simulated grid cells, using cover crops moderately enhanced N losses, in particular for N-fixing cover crops (F_{CC-BNF}, Fig. 5a). One possible explanation for this simulation is that the enhanced available N derived from the fast decomposition of cover crops would serve as substrate for N losses instead of being taken up by the main crop, mainly due to the temporal inconsistency between periods of soil N mineralization and high N demand of the main crop (Marcillo and Miguez, 2017). Compared to the bare fallow model setup (see F_{std}, Table 3), our simulations regionally predicted a slight yield increase of 6 % in non-legume cover crop systems but a high increase of 19 % for N-fixing cover crops (Fig. 5a), supporting the meta-analysis findings that legume cover crops usually contribute more to increasing

subsequent crop yields than non-legumes when N fertilizer inputs are low (Tonitto et al., 2006; Quemada et al., 2013; Marcillo and Miguez, 2017; Thapa et al., 2018). These slight yield benefits (also reduced productivity in few grid cells; see Fig. 5a) found in non-legume cover crop simulations primarily resulted from indirect competition for water and nutrients (Valkama et al., 2015), which may not be available for the following main crops planted in the long rainy season.

4.4 Trade-offs and win–win management options

In our study, we attempted to identify synergistic management strategies for achieving environmental sustainability without compromising crop production in eastern Africa. None of the assessed management options fully achieved a win–win situation in terms of increasing soil C stocks and crop production while minimizing N losses when integrated over the study regions. Synergies and trade-offs among the three examined indicators varied between locations (Fig. 5a) and cropping systems (Fig. S6).

From the perspectives of food demand and SOC sequestration only, conservation agriculture (CA, F_{conserv}), as an integrated management with no tillage, residue, and manure application, and N-fixing cover crops included, was simulated to be the most promising practice for both present-day conditions and future scenarios. Nevertheless, considering the potential yield reduction in the first several years under CA systems (Stevenson et al., 2014; Pittelkow et al., 2015b), it may be difficult to convince smallholder farmers to adopt such a practice; if indeed 1 %–25 % of crop production loss could be expected compared to the conventional management practice (Fig. 6c), farmers would suffer economic losses despite the accompanying 1 %–10 % of increase in SOC storage (Fig. 6a). Furthermore, labour demand and cost-ineffective investment in CA maintenance may prevent this practice from being implemented widely in eastern Africa (Thierfelder et al., 2013; Kihara et al., 2020). However, in our study this practice was modelled as the only one showing a net SOC sequestration in the future, with annual carbon uptake rates of 1.1 and 2.7 Tg C yr^{-1} between 2015 and 2100 for Kenya and Ethiopia, respectively (Fig. S8). The economic considerations would be potentially quite different if, in a carbon trading scheme, land management that leads to enhanced carbon sequestration would receive monetary compensation for the resulting yield reductions.

Rather than adopting a fully integrated CA, it is more common to use N-fixing legumes as cover crops or intercrops in smallholder farming systems over eastern Africa (Rao and Mathuva, 2000; Ngome et al., 2011). This crop management approach in our simulations ($F_{\text{CC-BNF}}$) also had positive impacts for soil C storage and food production, with low environmental cost in terms of N losses (Fig. 5). This win–win situation could also be sustained under future climate change (Fig. 6). However, it should be noted that the absence of soil pH constraints on legume inoculation in LPJ-GUESS

(Ma et al., 2022) most likely results in an overestimate of the N fixation rate in the $F_{\text{CC-BNF}}$ simulation. For example, our modelled N fixed by herbaceous legumes can be up to 70–90 kg N ha^{-1} yr^{-1} in some grid cells (Fig. S9 in the Supplement), while much experimental evidence from African farms indicates that the nodulation of roots in grain and forage legumes in SSA may not be successful, primarily caused by the inconsistent effectiveness of inoculation in the acid soils (Ulzen et al., 2016; Muleta et al., 2017; Vanlauwe et al., 2019). The overestimated N fixation in the model may thus bias the contribution of legume cover crops to the C–N cycle and crop production – but possibly also to N losses. To better represent cover crop management, the evaluation of modelled N-fixing herbaceous legumes against targeted field experiments, together with implementation of multi-cropping systems (see Sect.4.3.1), is needed in future model work.

5 Summary

In this study we presented a large-scale modelling analysis with LPJ-GUESS, highlighting potential long-term effects of management practices on crop ecosystems in eastern Africa under different climate change scenarios. The modelled C–N variables and crop yields in responses to varying agricultural practices were evaluated. Our results showed that crop ecosystems represented in LPJ-GUESS realistically responded to different management strategies and climate variation and produced soil C stocks, N losses, and crop productivity comparable to measurements in the studied region.

Our model demonstrated that the effects of management on agricultural ecosystems in eastern Africa can be beneficial for climate change mitigation without compromising crop yields, in particular for the combined conservation agriculture practice with all soil-C-conserving techniques included. This integrated strategy was the only practice simulated to potentially achieve a positive SOC sequestration under climate change. Adopting N-fixing cover crop systems was identified as a dominant practice to regionally increase food production and C storage in agricultural soils, with low environmental costs in the form of N losses. This win–win situation was shown to persist under a range of future climate pathways. However, processes missing from the model, such as multi-cropping system and N losses via runoff and soil erosion, might have biased our assessed management effects on crop ecosystems regionally.

The adoption of these management practices by farmers is promising from a climate change mitigation perspective but perhaps difficult to achieve in reality because of the yield losses in the first several years under conservation agriculture systems. Farmers are mostly risk-averse when faced with new management practices. To change this situation, a payment scheme for carbon sequestration legislated by the government or volunteered by corporations and individuals (Salzman et al., 2018) may be needed to fully compensate

for farmers' economic losses in eastern Africa, particularly in the context of future environmental change.

References

Abdalla, M., Hastings, A., Cheng, K., Yue, Q., Chadwick, D., Espenberg, M., Truu, J., Rees, R. M., and Smith, P.: A critical review of the impacts of cover crops on nitrogen leaching, net greenhouse gas balance and crop productivity, Glob. Change Biol., 25, 2530–2543, https://doi.org/10.1111/gcb.14644, 2019.

Adimassu, Z., Alemu, G., and Tamene, L.: Effects of tillage and crop residue management on runoff, soil loss and crop yield in the Humid Highlands of Ethiopia, Agr. Syst., 168, 11–18, https://doi.org/10.1016/j.agsy.2018.10.007, 2019.

Alemu, G. and Bayu, W.: Effects of farmyard manure and combined N and P fertilizer on sorghum and soil characteristics in northeastern Ethiopia, J. Sustain. Agr., 26, 23–41, https://doi.org/10.1300/J064v26n02_04, 2005.

Arneth, A., Olsson, L., Cowie, A., Erb, K. H., Hurlbert, M., Kurz, W. A., Mirzabaev, A., and Rounsevell, M. D. A.: Restoring Degraded Lands, Annu. Rev. Env. Resour., 46, 569–599, https://doi.org/10.1146/annurev-environ-012320-054809, 2021.

Ayuke, F. O., Pulleman, M. M., Vanlauwe, B., de Goede, R. G. M., Six, J., Csuzdi, C., and Brussaard, L.: Agricultural management affects earthworm and termite diversity across humid to semi-arid tropical zones, Agr. Ecosyst. Environ., 140, 148–154, https://doi.org/10.1016/j.agee.2010.11.021, 2011.

Badagliacca, G., Benítez, E., Amato, G., Badalucco, L., Giambalvo, D., Laudicina, V. A., and Ruisi, P.: Long-term no-tillage application increases soil organic carbon, nitrous oxide emissions and faba bean (*Vicia faba* L.) yields under rain-fed Mediterranean conditions, Sci. Total Environ., 639, 350–359, https://doi.org/10.1016/j.scitotenv.2018.05.157, 2018.

Baudron, F., Jaleta, M., Okitoi, O., and Tegegn, A.: Conservation agriculture in African mixed crop-livestock systems: Expanding the niche, Agr. Ecosyst. Environ., 187, 171–182, https://doi.org/10.1016/j.agee.2013.08.020, 2014.

Blanco-Canqui, H., Lal, R., Post, W. M., Izaurralde, R. C., and Shipitalo, M. J.: Soil hydraulic properties influenced by corn stover removal from no-till corn in Ohio, Soil Till. Res., 92, 144–155, https://doi.org/10.1016/j.still.2006.02.002, 2007.

Bondeau, A., Smith, P. C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-campen, H., Müller, C., Reichstein, M., and Smith, B.: Modelling the role of agriculture for the 20th century global terrestrial carbon balance, Glob. Change Biol., 13, 679–706, https://doi.org/10.1111/j.1365-2486.2006.01305.x, 2007.

Bouwman, L., Goldewijk, K. K., Van Der Hoek, K. W., Beusen, A. H. W., Van Vuuren, D. P., Willems, J., Rufino, M. C., and Stehfest, E.: Erratum: Exploring global changes in nitrogen and phosphorus cycles in agriculture induced by livestock production over the 1900–2050 period (Proceedings of the National Academy of Sciences of the United States of America (2013) https://doi.org/10.1073/pnas.1012878108, P. Natl. Acad. Sci. USA, 110, 21196, https://doi.org/10.1073/pnas.1206191109, 2013.

Carsky, R. J., Becker, M., and Hauser, S.: Mucuna Cover Crop Fallow Systems: Potential and limitations, in: Sustaining soil fertility in West Africa, edited by: Tian, G., Ishida, F., Keatinge, D., Carsky, R., and Wendt, J., SSSA Special Publication, Madison, USA, 111–135, ISBN 978-0-89118-838-4, 2001.

Central Statistical Agency: The federal democratic republic of Ethiopia central statistical agency, agricultural sample survey

2015/16 (2008 E.C), volume V: report on area, production and farm management practice of belg season crops for private peasant holdings, Addis Ababa, Ethiopia, 2016.

Chatskikh, D., Hansen, S., Olesen, J. E., and Petersen, B. M.: A simplified modelling approach for quantifying tillage effects on soil carbon stocks, Eur. J. Soil Sci., 60, 924–934, https://doi.org/10.1111/j.1365-2389.2009.01185.x, 2009.

Cheesman, S., Thierfelder, C., Eash, N. S., Kassie, G. T., and Frossard, E.: Soil carbon stocks in conservation agriculture systems of Southern Africa, Soil Till. Res., 156, 99–109, https://doi.org/10.1016/j.still.2015.09.018, 2016.

Corbeels, M., Graaff, J. De, Hycenth, T., Penot, E., Baudron, F., Naudin, K., Andrieu, N., Chirat, G., Schuler, J., Nyagumbo, I., Rusinamhodzi, L., Traore, K., Dulla, H., and Solomon, I.: Understanding the impact and adoption of conservation agriculture in Africa: A multi-scale analysis, Agr. Ecosyst. Environ., 187, 155–170, 2014.

Corbeels, M., Cardinael, R., Naudin, K., Guibert, H., and Torquebiau, E.: The 4 per 1000 goal and soil carbon storage under agroforestry and conservation agriculture systems in sub-Saharan Africa, Soil Till. Res., 188, 16–26, https://doi.org/10.1016/j.still.2018.02.015, 2019.

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H., and Buontempo, C.: WFDE5: bias-adjusted ERA5 reanalysis data for impact studies, Earth Syst. Sci. Data, 12, 2097–2120, https://doi.org/10.5194/essd-12-2097-2020, 2020.

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., and Hanasaki, N.: GSWP-2: Multimodel analysis and implications for our perception of the land surface, B. Am. Meteorol. Soc., 87, 1381–1397, https://doi.org/10.1175/BAMS-87-10-1381, 2006.

Elliott, J., Müller, C., Deryng, D., Chryssanthacopoulos, J., Boote, K. J., Büchner, M., Foster, I., Glotter, M., Heinke, J., Iizumi, T., Izaurralde, R. C., Mueller, N. D., Ray, D. K., Rosenzweig, C., Ruane, A. C., and Sheffield, J.: The Global Gridded Crop Model Intercomparison: data and modeling protocols for Phase 1 (v1.0), Geosci. Model Dev., 8, 261–277, https://doi.org/10.5194/gmd-8-261-2015, 2015.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

FAO: Food loss and food waste: Extent, Causes, and Prevention, Rome, Italy, 2011.

FAOSTAT: Production/Crops and livestock products, FAOSTAT [data set], https://www.fao.org/faostat/en/#data, last access: 19 August 2021.

Gichangi, E. M., Karanja, N. K., and Wood, C. W.: Composting cattle manure from zero grazing system with agro-organic wastes to minimise nitrogen losses in smallholder farms in kenya, Trop. Subtrop. Agroecosystems, 6, 57–64, 2006.

Giller, K. E., Witter, E., Corbeels, M., and Tittonell, P.: Conservation agriculture and smallholder farming in Africa: The heretics' view, Field Crop. Res., 114, 23–34, https://doi.org/10.1016/j.fcr.2009.06.017, 2009.

Githongo, M. W., Kiboi, M. N., Ngetich, F. K., Musafiri, C. M., Muriuki, A., and Fliessbach, A.: The effect of minimum tillage and animal manure on maize yields and soil organic

carbon in sub-Saharan Africa, Environ. Challenges, 5, 100340, https://doi.org/10.1016/j.envc.2021.100340, 2021.

Gross, A. and Glaser, B.: Meta-analysis on how manure application changes soil organic carbon storage, Sci. Rep.-UK, 11, 1–13, https://doi.org/10.1038/s41598-021-82739-7, 2021.

Haddaway, N. R., Hedlund, K., Jackson, L. E., Kätterer, T., Lugato, E., Thomsen, I. K., Jørgensen, H. B., and Isberg, P. E.: How does tillage intensity affect soil organic carbon? A systematic review, Environ. Evid., 6, 30, https://doi.org/10.1186/s13750-017-0108-9, 2017.

Harris, I., Osborn, T. J., Jones, P., and Lister, D.: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset, Sci. Data, 7, 109, https://doi.org/10.1038/s41597-020-0453-3, 2020 (data available at: https://data.ceda.ac.uk/badc/cru/data/cru_jra/cru_jra_2.1 (last access: 14 November 2021).

Herzfeld, T., Heinke, J., Rolinski, S., and Müller, C.: Soil organic carbon dynamics from agricultural management practices under climate change, Earth Syst. Dynam., 12, 1037–1055, https://doi.org/10.5194/esd-12-1037-2021, 2021.

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., van Vuuren, D. P., and Zhang, X.: Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6, Geosci. Model Dev., 13, 5425–5464, https://doi.org/10.5194/gmd-13-5425-2020, 2020.

Jobbágy, E. G. and Jackson, R. B.: The vertical distribution of soil organic carbon and its relation to climate and vegetation, Ecol. Appl., 10, 423–436, https://doi.org/10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2, 2000.

Johnston, A. E., Poulton, P. R., and Coleman, K.: Chapter 1 Soil Organic Matter. Its Importance in Sustainable Agriculture and Carbon Dioxide Fluxes, Adv. Agron., 101, 1–57, https://doi.org/10.1016/S0065-2113(08)00801-8, 2009.

Kaltenegger, K., Erb, K. H., Matej, S., and Winiwarter, W.: Gridded soil surface nitrogen surplus on grazing and agricultural land: Impact of land use maps, Environ. Res. Commun., 3, 055003, https://doi.org/10.1088/2515-7620/abedd8, 2021.

Kihara, J., Martius, C., Bationo, A., and Vlek, P. L. G.: Effects of tillage and crop residue application on Soybean nitrogen fixation in a tropical ferralsol, Agriculture, 1, 22–37, https://doi.org/10.3390/agriculture1010022, 2011.

Kihara, J., Martius, C., and Bationo, A.: Crop residue disappearance and macrofauna activity in sub-humid western Kenya, Nutr. Cycl. Agroecosys., 102, 101–111, https://doi.org/10.1007/s10705-014-9649-2, 2015.

Kihara, J., Bolo, P., Kinyua, M., Nyawira, S. S., and Sommer, R.: Soil health and ecosystem services: Lessons from sub-Sahara Africa (SSA), Geoderma, 370, 114342, https://doi.org/10.1016/j.geoderma.2020.114342, 2020.

Knapp, S. and van der Heijden, M. G. A.: A global meta-analysis of yield stability in organic and conservation agriculture, Nat. Commun., 9, 1–9, https://doi.org/10.1038/s41467-018-05956-1, 2018.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., and Kiyotoshi, T.: The JRA-55 reanalysis: General specifications and basic characteristics, J. Meteorol. Soc. Jpn., 93, 5–48, https://doi.org/10.2151/jmsj.2015-001, 2015 (data available at: https://data.ceda.ac.uk/badc/cru/data/cru_jra/cru_jra_2.1 (last access: 14 November 2021).

Laekemariam, F., Kibret, K., Mamo, T., Karltun, E., and Gebrekidan, H.: Physiographic characteristics of agricultural lands and farmers' soil fertility management practices in Wolaita zone, Southern Ethiopia, Environ. Syst. Res., 5, 24, https://doi.org/10.1186/s40068-016-0076-z, 2016.

Lal, R.: Soil carbon sequestration impacts on global climate change and food security, Science, 304, 1623–1627, https://doi.org/10.1126/science.1097396, 2004.

Lal, R.: Soil carbon stocks under present and future climate with specific reference to European ecoregions, Nutr. Cycl. Agroecosys., 81, 113–127, https://doi.org/10.1007/s10705-007-9147-x, 2008.

Lange, S.: Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0), Geosci. Model Dev., 12, 3055–3070, https://doi.org/10.5194/gmd-12-3055-2019, 2019.

Lange, S., Menz, C., Gleixner, S., Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Müller Schmied, H., Hersbach, H., Buontempo, C., and Cagnazzo, C.: WFDE5 over land merged with ERA5 over the ocean, Version 2.0, ISIMIP Repository [data set], https://doi.org/10.48364/ISIMIP.342217, 2021.

Lemma, B., Williams, S., and Paustian, K.: Long term soil carbon sequestration potential of smallholder croplands in southern Ethiopia with DAYCENT model, J. Environ. Manage., 294, 112893, https://doi.org/10.1016/j.jenvman.2021.112893, 2021.

Li, C., Frolking, S., and Harriss, R.: Modeling carbon biogeochemistry in agricultural soil, Global Biogeochem. Cy., 8, 237–254, 1994.

Lindeskog, M., Arneth, A., Bondeau, A., Waha, K., Seaquist, J., Olin, S., and Smith, B.: Implications of accounting for land use in simulations of ecosystem carbon cycling in Africa, Earth Syst. Dynam., 4, 385–407, https://doi.org/10.5194/esd-4-385-2013, 2013.

Liu, J., You, L., Amini, M., Obersteiner, M., Herrero, M., Zehnder, A. J. B., and Yang, H.: A high-resolution assessment on global nitrogen flows in cropland, P. Natl. Acad. Sci. USA, 107, 8035–8040, https://doi.org/10.1073/pnas.0913658107, 2010.

Liu, Y., Wu, L., Baddeley, J. A., and Watson, C. A.: Models of biological nitrogen fixation of legumes. A review, Agron. Sustain. Dev., 31, 155–172, https://doi.org/10.1051/agro/2010008, 2011.

Lugato, E., Bampa, F., Panagos, P., Montanarella, L., and Jones, A.: Potential carbon sequestration of European arable soils estimated by modelling a comprehensive set of management practices, Glob. Change Biol., 20, 3557–3567, https://doi.org/10.1111/gcb.12551, 2015.

Ma, J., Olin, S., Anthoni, P., Rabin, S. S., Bayer, A. D., Nyawira, S. S., and Arneth, A.: Modeling symbiotic biological nitrogen fixation in grain legumes globally with LPJ-GUESS (v4.0, r10285), Geosci. Model Dev., 15, 815–839, https://doi.org/10.5194/gmd-15-815-2022, 2022.

Marcillo, G. S. and Miguez, F. E.: Corn yield response to winter cover crops: An updated meta-analysis, J. Soil Water Conserv., 72, 226–239, https://doi.org/10.2489/jswc.72.3.226, 2017.

Martinsen, V., Munera-Echeverri, J. L., Obia, A., Cornelissen, G., and Mulder, J.: Significant build-up of soil organic carbon under climate-smart conservation farming in Sub-Saharan Acrisols, Sci. Total Environ., 660, 97–104, https://doi.org/10.1016/j.scitotenv.2018.12.452, 2019.

Meinshausen, M., Nicholls, Z. R. J., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., Canadell, J. G., Daniel, J. S., John, A., Krummel, P. B., Luderer, G., Meinshausen, N., Montzka, S. A., Rayner, P. J., Reimann, S., Smith, S. J., van den Berg, M., Velders, G. J. M., Vollmer, M. K., and Wang, R. H. J.: The shared socioeconomic pathway (SSP) greenhouse gas concentrations and their extensions to 2500, Geosci. Model Dev., 13, 3571–3605, https://doi.org/10.5194/gmd-13-3571-2020, 2020.

Mesfin, S., Gebresamuel, G., Haile, M., and Zenebe, A.: Modelling spatial and temporal soil organic carbon dynamics under climate and land management change scenarios, northern Ethiopia, Eur. J. Soil Sci., 72, 1298–1311, https://doi.org/10.1111/ejss.13060, 2021.

Moebius-Clune, B. N., van Es, H. M., Idowu, O. J., Schindelbeck, R. R., Kimetu, J. M., Ngoze, S., Lehmann, J., and Kinyangi, J. M.: Long-term soil quality degradation along a cultivation chronosequence in western Kenya, Agr. Ecosyst. Environ., 141, 86–99, https://doi.org/10.1016/j.agee.2011.02.018, 2011.

Mugizi, F. M. P. and Matsumoto, T.: Population pressure and soil quality in Sub-Saharan Africa: Panel evidence from Kenya, Land Use Policy, 94, 104499, https://doi.org/10.1016/j.landusepol.2020.104499, 2020.

Muleta, D., Ryder, M. H., and Denton, M. D.: The potential for rhizobial inoculation to increase soybean grain yields on acid soils in Ethiopia, Soil Sci. Plant Nutr., 63, 441–451, https://doi.org/10.1080/00380768.2017.1370961, 2017.

Muoni, T., Barnes, A. P., Öborn, I., Watson, C. A., Bergkvist, G., Shiluli, M., and Duncan, A. J.: Farmer perceptions of legumes and their functions in smallholder farming systems in east Africa, Int. J. Agr. Sustain., 17, 205–218, https://doi.org/10.1080/14735903.2019.1609166, 2019.

Ngome, A. F., Becker, M., Mtei, K. M., and Mussgnug, F.: Fertility management for maize cultivation in some soils of Western Kenya, Soil Till. Res., 117, 69–75, https://doi.org/10.1016/j.still.2011.08.010, 2011.

Nyawira, S. S., Hartman, M. D., Nguyen, T. H., Margenot, A. J., Kihara, J., Paul, B. K., Williams, S., Bolo, P., and Sommer, R.: Simulating soil organic carbon in maize-based systems under improved agronomic management in Western Kenya, Soil Till. Res., 211, 105000, https://doi.org/10.1016/j.still.2021.105000, 2021.

Okosun, O. O., Allen, K. C., Glover, J. P., and Reddy, G. V. P.: Biology, Ecology, and Management of Key Sorghum Insect Pests, J. Integr. Pest Manag., 12, 4, https://doi.org/10.1093/jipm/pmaa027, 2021.

Olin, S., Schurgers, G., Lindeskog, M., Wårlind, D., Smith, B., Bodin, P., Holmér, J., and Arneth, A.: Modelling the response of yields and tissue C : N to changes in atmospheric CO_2 and N management in the main wheat regions of western Europe, Biogeosciences, 12, 2489–2515, https://doi.org/10.5194/bg-12-2489-2015, 2015a.

Olin, S., Lindeskog, M., Pugh, T. A. M., Schurgers, G., Wårlind, D., Mishurov, M., Zaehle, S., Stocker, B. D., Smith, B., and Ar-

neth, A.: Soil carbon management in large-scale Earth system modelling: implications for crop yields and nitrogen leaching, Earth Syst. Dynam., 6, 745–768, https://doi.org/10.5194/esd-6-745-2015, 2015b.

Olsson, L., Barbosa, H., Bhadwal, S., Cowie, A., Delusca, K., Flores-Renteria, D., Hermans, K., Jobbagy, E., Kurz, W., Li, D., Sonwa, D. J., and Stringer, L.: Land Degradation, in: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems, edited by: Shukla, P. R., Skea, J., Calvo Buendia, E., Masson-Delmotte, V., Pörtner, H. O., Roberts, D. C., Zhai, P., Slade, R., Connors, S., van Diemen, R., Ferrat, M., Haughey, E., Luz, S., Neogi, S., Pathak, M., Petzold, J., Portugal Pereira, J., Vyas, P., Huntley, E., Kissick, K., Belkacemi, M., and Malley, J., IPCC, Geneva, Switzerland, pp. 345–436, 2019.

Parton, W. J., Scurlock, J. M. O., Ojima, D. S., Gilmanov, T. G., Scholes, R. J., Schimel, D. S., Kirchner, T., Menaut, J. -C, Seastedt, T., Garcia Moya, E., Kamnalrut, A., and Kinyamario, J. I.: Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide, Global Biogeochem. Cy., 7, 785–809, https://doi.org/10.1029/93GB02042, 1993.

Penning de Vries, F. W. T., Jansen, D. M., Berge, H. F. M. ten, and Bakema, A.: Simulation of ecophysiological processes of growth in several annual crops, Centre for Agricultural Publishing and Documentation (Pudoc), Wageningen, the Netherlands, ISBN 90-220-1000-7, 1989.

Pittelkow, C. M., Liang, X., Linquist, B. A., Van Groenigen, L. J., Lee, J., Lundy, M. E., Van Gestel, N., Six, J., Venterea, R. T., and Van Kessel, C.: Productivity limits and potentials of the principles of conservation agriculture, Nature, 517, 365–368, https://doi.org/10.1038/nature13809, 2015a.

Pittelkow, C. M., Linquist, B. A., Lundy, M. E., Liang, X., van Groenigen, K. J., Lee, J., van Gestel, N., Six, J., Venterea, R. T., and van Kessel, C.: When does no-till yield more? A global meta-analysis, Field Crop. Res., 183, 156–168, https://doi.org/10.1016/j.fcr.2015.07.020, 2015b.

Poeplau, C. and Don, A.: Carbon sequestration in agricultural soils via cultivation of cover crops – A meta-analysis, Agr. Ecosyst. Environ., 200, 33–41, https://doi.org/10.1016/j.agee.2014.10.024, 2015.

Portmann, F. T., Siebert, S., and Döll, P.: MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling, Global Biogeochem. Cy., 24, 1–24, https://doi.org/10.1029/2008gb003435, 2010.

Porwollik, V., Rolinski, S., Heinke, J., von Bloh, W., Schaphoff, S., and Müller, C.: The role of cover crops for cropland soil carbon, nitrogen leaching, and agricultural yields – a global simulation study with LPJmL (V. 5.0-tillage-cc), Biogeosciences, 19, 957–977, https://doi.org/10.5194/bg-19-957-2022, 2022.

Powlson, D. S., Stirling, C. M., Thierfelder, C., White, R. P., and Jat, M. L.: Does conservation agriculture deliver climate change mitigation through soil carbon sequestration in tropical agro-ecosystems?, Agr. Ecosyst. Environ., 220, 164–174, https://doi.org/10.1016/j.agee.2016.01.005, 2016.

Pugh, T. A. M., Arneth, A., Olin, S., Ahlström, A., Bayer, A. D., Klein Goldewijk, K., Lindeskog, M., and Schurgers, G.: Simulated carbon emissions from land-use change are substantially enhanced by accounting for agricultural management, Environ. Res. Lett., 10, https://doi.org/10.1088/1748-9326/10/12/124008, 2015.

Quemada, M., Baranski, M., Nobel-de Lange, M. N. J., Vallejo, A., and Cooper, J. M.: Meta-analysis of strategies to control nitrate leaching in irrigated agricultural systems and their effects on crop yield, Agr. Ecosyst. Environ., 174, 1–10, https://doi.org/10.1016/j.agee.2013.04.018, 2013.

Rao, M. R. and Mathuva, M. N.: Legumes for improving maize yields and income in semi-arid Kenya, Agr. Ecosyst. Environ., 78, 123–137, https://doi.org/10.1016/S0167-8809(99)00125-5, 2000.

Reay, D. S., Davidson, E. A., Smith, K. A., Smith, P., Melillo, J. M., Dentener, F., and Crutzen, P. J.: Global agriculture and nitrous oxide emissions, Nat. Clim. Change, 2, 410–416, https://doi.org/10.1038/nclimate1458, 2012.

Ruddiman, W. F.: The Anthropogenic Greenhouse Era Began Thousands of Years Ago, Climatic Change, 61, 261–293, 2003.

Salzman, J., Bennett, G., Carroll, N., Goldstein, A., and Jenkins, M.: The global status and trends of Payments for Ecosystem Services, Nat. Sustain., 1, 136–144, https://doi.org/10.1038/s41893-018-0033-0, 2018.

Sanderman, J., Hengl, T., and Fiske, G. J.: Soil carbon debt of 12,000 years of human land use, P. Natl. Acad. Sci. USA, 114, 9575–9580, https://doi.org/10.1073/pnas.1800925115, 2017.

Seufert, V., Ramankutty, N., and Foley, J. A.: Comparing the yields of organic and conventional agriculture, Nature, 485, 229–232, https://doi.org/10.1038/nature11069, 2012.

Shang, Z., Abdalla, M., Xia, L., Zhou, F., Sun, W., and Smith, P.: Can cropland management practices lower net greenhouse emissions without compromising yield?, Glob. Change Biol., 27, 4657–4670, https://doi.org/10.1111/gcb.15796, 2021.

Sheahan, M. and Barrett, C. B.: Food loss and waste in Sub-Saharan Africa: A critical review, Food Policy, 70, 1–12, https://doi.org/10.1016/j.foodpol.2017.03.012, 2017.

Sileshi, G., Akinnifesi, F. K., Ajayi, O. C., and Place, F.: Meta-analysis of maize yield response to woody and herbaceous legumes in sub-Saharan Africa, Plant Soil, 307, 1–19, https://doi.org/10.1007/s11104-008-9547-y, 2008.

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., and Zaehle, S.: Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model, Biogeosciences, 11, 2027–2054, https://doi.org/10.5194/bg-11-2027-2014, 2014.

Smith, P., House, J. I., Bustamante, M., Sobocká, J., Harper, R., Pan, G., West, P. C., Clark, J. M., Adhya, T., Rumpel, C., Paustian, K., Kuikman, P., Cotrufo, M. F., Elliott, J. A., Mcdowell, R., Griffiths, R. I., Asakawa, S., Bondeau, A., Jain, A. K., Meersmans, J., and Pugh, T. A. M.: Global change pressures on soils from land use and management, Glob. Change Biol., 22, 1008–1028, https://doi.org/10.1111/gcb.13068, 2016.

Smith, P., Calvin, K., Nkem, J., Campbell, D., Cherubini, F., Grassi, G., Korotkov, V., Le Hoang, A., Lwasa, S., McElwee, P., Nkonya, E., Saigusa, N., Soussana, J. F., Taboada, M. A., Manning, F. C., Nampanzira, D., Arias-Navarro, C., Vizzarri, M., House, J., Roe, S., Cowie, A., Rounsevell, M., and Arneth, A.: Which practices co-deliver food security, climate change mitigation and adaptation, and combat land degrada-

tion and desertification?, Glob. Change Biol., 26, 1532–1575, https://doi.org/10.1111/gcb.14878, 2020.

Smith, W. N., Grant, B. B., Campbell, C. A., McConkey, B. G., Desjardins, R. L., Kröbel, R., and Malhi, S. S.: Crop residue removal effects on soil carbon: Measured and inter-model comparisons, Agr. Ecosyst. Environ., 161, 27–38, https://doi.org/10.1016/j.agee.2012.07.024, 2012.

Sommer, R. and Bossio, D.: Dynamics and climate change mitigation potential of soil organic carbon sequestration, J. Environ. Manage., 144, 83–87, https://doi.org/10.1016/j.jenvman.2014.05.017, 2014.

Sommer, R., Paul, B. K., Mukalama, J., and Kihara, J.: Reducing losses but failing to sequester carbon in soils – the case of Conservation Agriculture and Integrated Soil Fertility Management in the humid tropical agro-ecosystem of Western Kenya, Agr. Ecosyst. Environ., 254, 82–91, https://doi.org/10.1016/j.agee.2017.11.004, 2018.

Stathers, T., Holcroft, D., Kitinoja, L., Mvumi, B. M., English, A., Omotilewa, O., Kocher, M., Ault, J., and Torero, M.: A scoping review of interventions for crop postharvest loss reduction in sub-Saharan Africa and South Asia, Nat. Sustain., 3, 821–835, https://doi.org/10.1038/s41893-020-00622-1, 2020.

Stevenson, J. R., Serraj, R., and Cassman, K. G.: Evaluating conservation agriculture for small-scale farmers in Sub-Saharan Africa and South Asia, Agr. Ecosyst. Environ., 187, 1–10, https://doi.org/10.1016/j.agee.2014.01.018, 2014.

Thapa, R., Mirsky, S. B., and Tully, K. L.: Cover Crops Reduce Nitrate Leaching in Agroecosystems:A Global Meta-Analysis, J. Environ. Qual., 47, 1400–1411, https://doi.org/10.2134/jeq2018.03.0107, 2018.

Thierfelder, C., Mwila, M., and Rusinamhodzi, L.: Conservation agriculture in eastern and southern provinces of Zambia: Long-term effects on soil quality and maize productivity, Soil Till. Res., 126, 246–258, https://doi.org/10.1016/j.still.2012.09.002, 2013.

Tian, H., Yang, J., Lu, C., Xu, R., Canadell, J. G., Jackson, R. B., Arneth, A., Chang, J., Chen, G., Ciais, P., Gerber, S., Ito, A., Huang, Y., Joos, F., Lienert, S., Messina, P., Olin, S., Pan, S., Peng, C., Saikawa, E., Thompson, R. L., Vuichard, N., Winiwarter, W., Zaehle, S., Zhang, B., Zhang, K., and Zhu, Q.: The global N_2O model intercomparison project, B. Am. Meteorol. Soc., 99, 1231–1251, https://doi.org/10.1175/BAMS-D-17-0212.1, 2018.

Tian, H., Xu, R., Canadell, J. G., Thompson, R. L., Winiwarter, W., Suntharalingam, P., Davidson, E. A., Ciais, P., Jackson, R. B., Janssens-Maenhout, G., Prather, M. J., Regnier, P., Pan, N., Pan, S., Peters, G. P., Shi, H., Tubiello, F. N., Zaehle, S., Zhou, F., Arneth, A., Battaglia, G., Berthet, S., Bopp, L., Bouwman, A. F., Buitenhuis, E. T., Chang, J., Chipperfield, M. P., Dangal, S. R. S., Dlugokencky, E., Elkins, J. W., Eyre, B. D., Fu, B., Hall, B., Ito, A., Joos, F., Krummel, P. B., Landolfi, A., Laruelle, G. G., Lauerwald, R., Li, W., Lienert, S., Maavara, T., MacLeod, M., Millet, D. B., Olin, S., Patra, P. K., Prinn, R. G., Raymond, P. A., Ruiz, D. J., van der Werf, G. R., Vuichard, N., Wang, J., Weiss, R. F., Wells, K. C., Wilson, C., Yang, J., and Yao, Y.: A comprehensive quantification of global nitrous oxide sources and sinks, Nature, 586, 248–256, https://doi.org/10.1038/s41586-020-2780-0, 2020.

Tonitto, C., David, M. B., and Drinkwater, L. E.: Replacing bare fallows with cover crops in fertilizer-intensive cropping systems: A meta-analysis of crop yield and N dynamics, Agr. Ecosyst. Environ., 112, 58–72, https://doi.org/10.1016/j.agee.2005.07.003, 2006.

Ulzen, J., Abaidoo, R. C., Mensah, N. E., Masso, C., and Abdel-Gadir, A. A. H.: Bradyrhizobium inoculants enhance grain yields of soybean and cowpea in Northern Ghana, Front. Plant Sci., 7, 1–9, https://doi.org/10.3389/fpls.2016.01770, 2016.

Valbuena, D., Erenstein, O., Homann-Kee Tui, S., Abdoulaye, T., Claessens, L., Duncan, A. J., Gérard, B., Rufino, M. C., Teufel, N., van Rooyen, A., and van Wijk, M. T.: Conservation Agriculture in mixed crop-livestock systems: Scoping crop residue trade-offs in Sub-Saharan Africa and South Asia, Field Crop. Res., 132, 175–184, https://doi.org/10.1016/j.fcr.2012.02.022, 2012.

Valkama, E., Lemola, R., Känkänen, H., and Turtola, E.: Meta-analysis of the effects of undersown catch crops on nitrogen leaching loss and grain yields in the Nordic countries, Agr. Ecosyst. Environ., 203, 93–101, https://doi.org/10.1016/j.agee.2015.01.023, 2015.

Vanlauwe, B., Hungria, M., Kanampiu, F., and Giller, K. E.: The role of legumes in the sustainable intensification of African smallholder agriculture: Lessons learnt and challenges for the future, Agr. Ecosyst. Environ., 284, 106583, https://doi.org/10.1016/j.agee.2019.106583, 2019.

Volkholz, J. and Müller, C.: ISIMIP3 soil input data (v1.0), ISIMIP Repos. [data set], https://doi.org/10.48364/ISIMIP.942125, 2020.

Waha, K., Van Bussel, L. G. J., Müller, C., and Bondeau, A.: Climate-driven simulation of global crop sowing dates, Global Ecol. Biogeogr., 21, 247–259, https://doi.org/10.1111/j.1466-8238.2011.00678.x, 2012.

Wainwright, C. M., Marsham, J. H., Keane, R. J., Rowell, D. P., Finney, D. L., Black, E., and Allan, R. P.: 'Eastern African Paradox' rainfall decline due to shorter not less intense Long Rains, npj Clim. Atmos. Sci., 2, 1–9, https://doi.org/10.1038/s41612-019-0091-7, 2019.

Wårlind, D., Smith, B., Hickler, T., and Arneth, A.: Nitrogen feedbacks increase future terrestrial ecosystem carbon uptake in an individual-based dynamic vegetation model, Biogeosciences, 11, 6131–6146, https://doi.org/10.5194/bg-11-6131-2014, 2014.

Wynants, M., Kelly, C., Mtei, K., Munishi, L., Patrick, A., Rabinovich, A., Nasseri, M., Gilvear, D., Roberts, N., Boeckx, P., Wilson, G., Blake, W. H., and Ndakidemi, P.: Drivers of increased soil erosion in East Africa's agro-pastoral systems: changing interactions between the social, economic and natural domains, Reg. Environ. Change, 19, 1909–1921, https://doi.org/10.1007/s10113-019-01520-9, 2019.

Xia, L., Lam, S. K., Wolf, B., Kiese, R., Chen, D., and Butterbach-Bahl, K.: Trade-offs between soil carbon sequestration and reactive nitrogen losses under straw return in global agroecosystems, Glob. Change Biol., 24, 5919–5932, https://doi.org/10.1111/gcb.14466, 2018.

Zaehle, S. and Friend, A. D.: Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates, Global Biogeochem. Cy., 24, 1–13, https://doi.org/10.1029/2009GB003521, 2010.

Zhang, B., Tian, H., Lu, C., Dangal, S. R. S., Yang, J., and Pan, S.: Global manure nitrogen production and application in cropland during 1860–2014: a 5 arcmin gridded global dataset

for Earth system modeling, Earth Syst. Sci. Data, 9, 667–678, https://doi.org/10.5194/essd-9-667-2017, 2017.

Zhang, X., Zou, T., Lassaletta, L., Mueller, N. D., Tubiello, F. N., Lisk, M. D., Lu, C., Conant, R. T., Dorich, C. D., Gerber, J., Tian, H., Bruulsema, T., Maaz, T. M. C., Nishina, K., Bodirsky, B. L., Popp, A., Bouwman, L., Beusen, A., Chang, J., Havlík, P., Leclère, D., Canadell, J. G., Jackson, R. B., Heffer, P., Wanner, N., Zhang, W., and Davidson, E. A.: Quantification of global and national nitrogen budgets for crop production, Nat. Food, 2, 529–540, https://doi.org/10.1038/s43016-021-00318-5, 2021.

Zhu, Y., Merbold, L., Leitner, S., Pelster, D. E., Okoma, S. A., Ngetich, F., Onyango, A. A., Pellikka, P., and Butterbach-Bahl, K.: The effects of climate on decomposition of cattle, sheep and goat manure in Kenyan tropical pastures, Plant Soil, 451, 325–343, https://doi.org/10.1007/s11104-020-04528-x, 2020.

Zomer, R. J., Bossio, D. A., Sommer, R., and Verchot, L. V.: Global Sequestration Potential of Increased Organic Carbon in Cropland Soils, Sci. Rep.-UK, 7, 1–8, https://doi.org/10.1038/s41598-017-15794-8, 2017.