Global mapping of crop-specific emission factors highlights hotspots of nitrous oxide mitigation

Article *in* Nature Food · November 2021

DOI: 10.1038/s43016-021-00384-9

19 authors, including:

Xiaoqing Cui
Peking University
8 PUBLICATIONS **166** CITATIONS
SEE PROFILE

Feng Zhou
University of Kashmir
86 PUBLICATIONS **3,194** CITATIONS
SEE PROFILE

Francesco Nicola Tubiello
Food and Agriculture Organization of the United Nations
219 PUBLICATIONS **21,333** CITATIONS
SEE PROFILE

Nathaniel D. Mueller
University of California, Irvine
94 PUBLICATIONS **15,777** CITATIONS
SEE PROFILE

nature food

ARTICLES

https://doi.org/10.1038/s43016-021-00384-9

Global mapping of crop-specific emission factors highlights hotspots of nitrous oxide mitigation

Xiaoqing Cui[1], Feng Zhou[1]✉, Philippe Ciais[2,3], Eric A. Davidson[4], Francesco N. Tubiello[5], Xiaoyue Niu[6], Xiaotang Ju[7], Josep G. Canadell[8], Alexander F. Bouwman[9,10,11], Robert B. Jackson[12], Nathaniel D. Mueller[13], Xunhua Zheng[14], David R. Kanter[15], Hanqin Tian[16], Wulahati Adalibieke[1], Yan Bo[1], Qihui Wang[1], Xiaoying Zhan[17] and Dongqiang Zhu[1]

Mitigating soil nitrous oxide (N_2O) emissions is essential for staying below a 2 °C warming threshold. However, accurate assessments of mitigation potential are limited by uncertainty and variability in direct emission factors (EFs). To assess where and why EFs differ, we created high-resolution maps of crop-specific EFs based on 1,507 georeferenced field observations. Here, using a data-driven approach, we show that EFs vary by two orders of magnitude over space. At global and regional scales, such variation is primarily driven by climatic and edaphic factors rather than the well-recognized management practices. Combining spatially explicit EFs with N surplus information, we conclude that global mitigation potential without compromising crop production is 30% (95% confidence interval, 17–53%) of direct soil emissions of N_2O, equivalent to the entire direct soil emissions of China and the United States combined. Two-thirds (65%) of the mitigation potential could be achieved on one-fifth of the global harvested area, mainly located in humid subtropical climates and across gleysols and acrisols. These findings highlight the value of a targeted policy approach on global hotspots that could deliver large N_2O mitigation as well as environmental and food co-benefits.

Nitrogen (N) is a key plant micronutrient applied in fertilizers to increase crop yield[1]. However, N applications to agricultural soils have become the largest anthropogenic source of nitrous oxide (N_2O) emissions globally, regionally and at the country scale[2,3]. Reducing agricultural N_2O emissions is thus conducive to achieving low warming targets[4] and preventing stratospheric ozone depletion[5], if the high-level crop yield achieved in recent decades is maintained to ensure future food security[6]. Yet the magnitude and pattern of the achievable N_2O mitigation potential that would not compromise crop yield remain poorly known. This is a barrier for nations and international organizations in developing and implementing effective N_2O mitigation programmes.

The mitigation potential of direct soil emissions of N_2O depends on the extent and location of N input reductions and the local values of direct emission factors (EFs, defined as the percentage of applied N emitted directly as N_2O-N). Although many estimates exist of the N input reductions necessary to stay within planetary boundaries[7–9], a global, spatially explicit quantification of EFs that considers the combined effects of environmental and management-related variables (for example, climate, soil, fertilization, irrigation and tillage) is lacking[10–12]. The key reason for this is that most field observations do not report the full diversity of EFs across soil–crop systems. Data-deficient countries thus apply Tier 1 default EF values from the Intergovernmental Panel on Climate Change (IPCC) guidelines[13] when reporting their greenhouse gas emissions to the United Nations Climate Convention. Such EFs are spatially and temporally constant, while recent studies have updated EFs by integrating the nonlinear response of N_2O emissions to agricultural management practices such as N application rate[14,15] or other management-related variables[16]. However, the paucity of globally representative observation data used for modelling EFs would result in inaccurate extrapolation when expanded to global predictions[17]. Moreover, these estimates do not account for finer-scale variation due to local environmental and management-related conditions.

Here, we take one step forwards by using a data-driven approach connecting crop-specific EF variations to climate, soil, fertilization, irrigation and tillage, on the basis of an extensive compilation of 1,507 chamber-based field observations of EFs spanning 234 sites and 31 countries from 1980 to 2018 (Supplementary Fig. 1 and Supplementary Data 1). With this dataset, we address three key

[1]Sino-France Institute of Earth Systems Science, Laboratory for Earth Surface Processes, College of Urban and Environmental Sciences, Peking University, Beijing, China. [2]Laboratoire des Sciences du Climat et de l'Environnement, LSCE, Gif sur Yvette, France. [3]Climate and Atmosphere Research Center (CARE-C), The Cyprus Institute, Nicosia, Cyprus. [4]Appalachian Laboratory, University of Maryland Center for Environmental Science, Frostburg, MD, USA. [5]Statistics Division, Food and Agriculture Organization of the United Nations, Rome, Italy. [6]Department of Statistics, The Pennsylvania State University, State College, PA, USA. [7]College of Tropical Crops, Hainan University, Haikou, China. [8]Global Carbon Project, CSIRO Oceans and Atmosphere, Canberra, Australian Capital Territory, Australia. [9]Department of Earth Sciences – Geochemistry, Faculty of Geosciences, Utrecht University, Utrecht, the Netherlands. [10]PBL Netherlands Environmental Assessment Agency, the Hague, the Netherlands. [11]Key Laboratory of Marine Chemistry Theory and Technology, Ministry of Education, Ocean University of China, Qingdao, China. [12]Department of Earth System Science, Woods Institute for the Environment, and Precourt Institute for Energy, Stanford University, Stanford, CA, USA. [13]Department of Ecosystem Science and Sustainability and Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO, USA. [14]State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China. [15]Department of Environmental Studies, New York University, New York, NY, USA. [16]International Center for Climate and Global Change Research, School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL, USA. [17]Agricultural Clean Watershed Research Group, Institute of Environment and Sustainable Development in Agriculture, Chinese Academy of Agricultural Sciences, Beijing, China. ✉e-mail: zhouf@pku.edu.cn

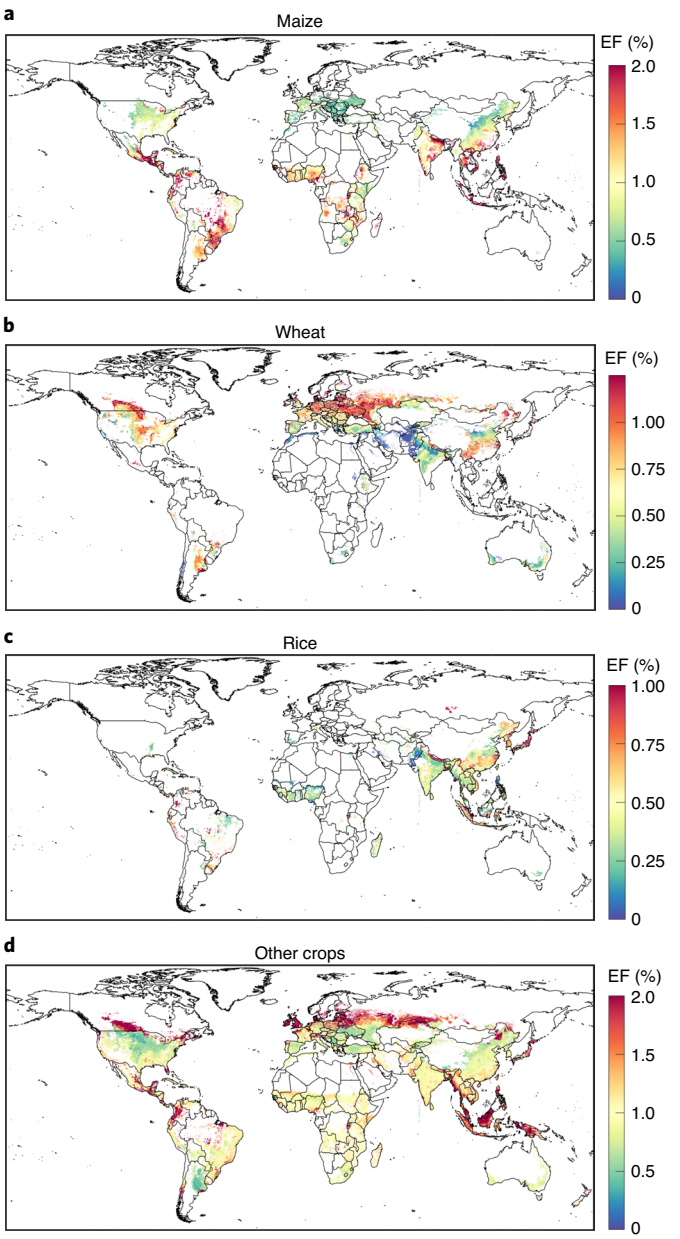

Fig. 1 | Spatial patterns of N_2O EFs for direct soil emissions. a, Maize. **b**, Wheat. **c**, Rice. **d**, Other crops. EFs are predicted with LME models, with the model uncertainty as illustrated in Supplementary Fig. 3. Values are shown only where the proportion of harvested area within the grid cell is greater than 0.5%. The map was generated in MATLAB R2020a (MATLAB and Statistics Toolbox Release R2020a, the MathWorks). The base map of the country boundaries was from the Global Administrative Areas dataset (https://gadm.org/download_world.html).

questions. First, what is the heterogeneity of EFs across global cropland? Second, how strongly do different drivers influence variations in EFs at global and regional scales? Third, how much global mitigation could be achieved while maintaining crop yield? We focus on croplands encompassing arable land and permanent crops, which dominate global fertilizer N applications (>95%)[18].

Our analysis expands on previous research in at least three aspects. First, we used a twofold to fourfold larger observation dataset of EFs than refs. [15,19,20] (Supplementary Text 1 and Supplementary Table 1), with driver values in field sites spanning >80% of the full covariate space across global cropland (Supplementary Fig. 2). We also mapped

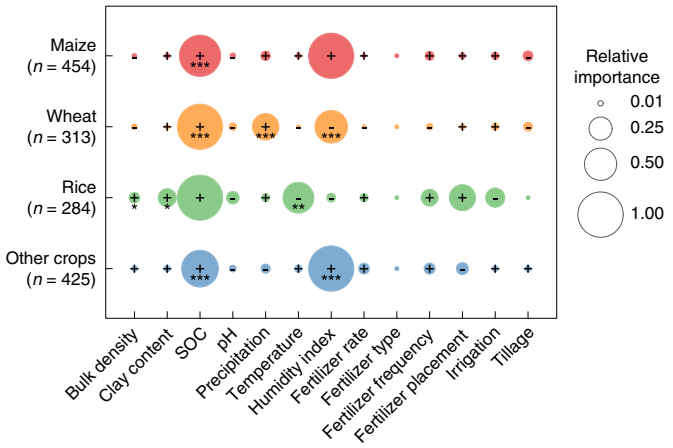

Fig. 2 | Relative importance by variable in shaping EF patterns. The rows show the results for each crop. The columns represent the variables that are included in the multi-model inference. In each row, the circle size of the variable groups is proportional to the relative change in importance. The circle size should be compared only within a row. The asterisks indicate the statistical significance of the effect (*$P<0.05$; ** $P<0.01$; ***$P<0.001$). The symbols + and − indicate positive and negative effects of the variables on N_2O EFs, respectively. Each management-related variable from fertilizer rate to tillage indicates the effect relative to zero N application, urea type, multiple applications, deep placement, rainfed (or intermittently flooding for rice) or no-till, respectively. No statistical significance or symbols are labelled for 'fertilizer type', as there is a combination of effects from multiple fertilizer types.

crop-specific EFs at five-arcminute spatial resolution using linear mixed-effects (LME) models constrained by the global observation dataset, avoiding assumptions associated with the representation of complex N_2O production processes as in process-based models[3]. Finally, we identified the global hotspots of cropland N_2O mitigation potential by crossing information on EFs and N input reduction data. Our results are timely given the current growth of N_2O emissions exceeding the highest projected emission scenario[3], and they bring a paradigm shift from asking, 'How much of global emissions can we mitigate?' to the more policy-relevant question of 'Where are the best opportunities to mitigate emissions most effectively?'.

Results and discussion

Pattern of crop-specific EFs. The global EF was estimated at 1.02% with a 95% confidence interval (CI) from 0.32% to 2.54% for maize due to the uncertainties stemming from sampling, modelling and input data, 0.58% (0.10–1.60%) for wheat, 0.52% (0.15–1.31%) for rice, and 1.20% (0.31–2.29%) for other crops. These EF means and CIs are in broad agreement with IPCC Tier 1 default values[13] (Supplementary Table 2), except for wheat, where our estimation is significantly smaller.

The spatial heterogeneity of N_2O EFs was striking regardless of crop type (Fig. 1) and despite the largest uncertainty of EFs being found in high-latitude areas for maize and in the tropics for wheat and rice (Supplementary Fig. 3). Areas with either high or low EFs are found on all continents, with EFs ranging from 0.08% to 3.77% for maize across the globe, 0.03–2.42% for wheat, 0.03–1.90% for rice and 0.01–3.12% for other crops. Areas with high EF values where observations were numerous were well captured by our global prediction maps (Fig. 1). Areas with larger EFs where few observations existed, including northern Europe, Central America, Southeast Asia and northern parts of South America, were also well constrained because of the high interpolation capability of our models (Supplementary Fig. 2).

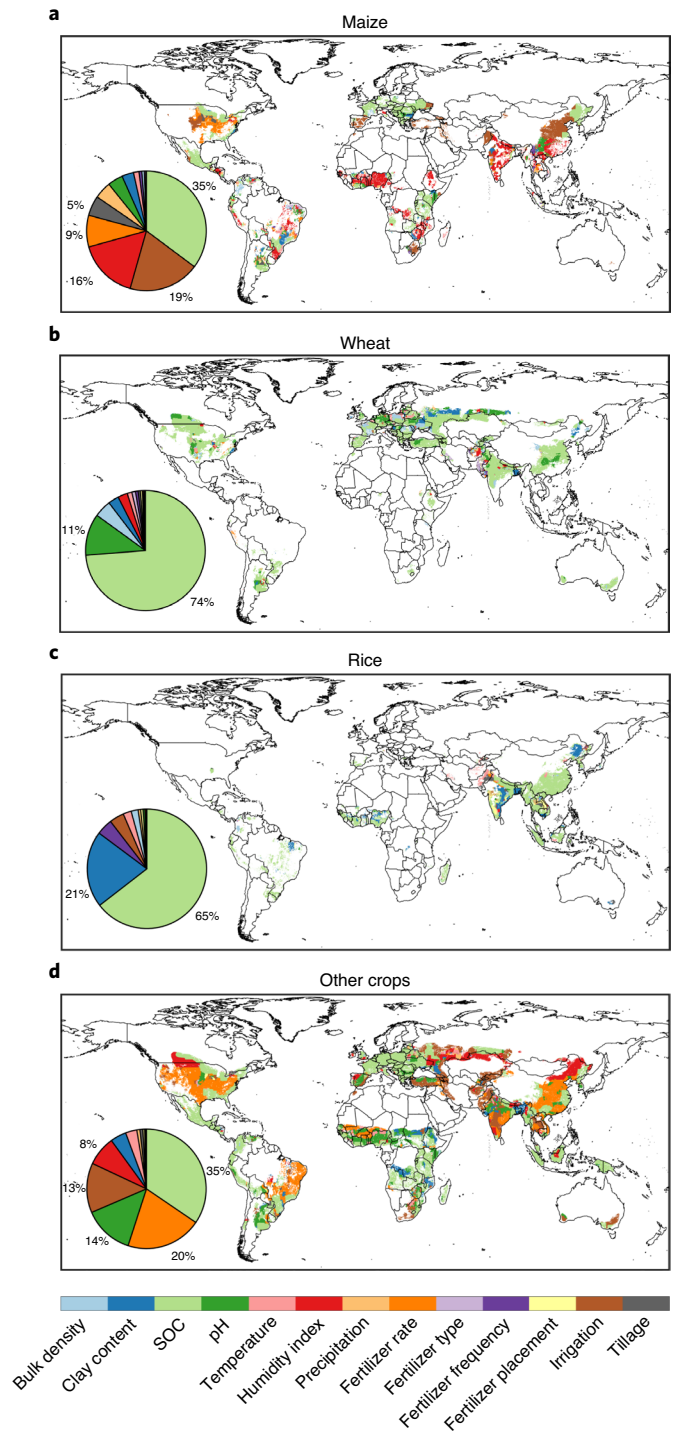

Fig. 3 | Distribution of dominant drivers regulating variation in N_2O EFs. a, Maize. **b**, Wheat. **c**, Rice. **d**, Other crops. The dominant driver is defined as the factor with the largest absolute value of the partial correlation coefficient (r) in each grid cell, where r between EFs and predictors is done for 3.75°-by-3.75° moving windows. Significant correlations ($P < 0.05$) are shown. The inset pie plots represent the ratio (%) of harvested areas for which EF variation is regulated by the dominant drivers. The map was generated in MATLAB R2020a (MATLAB and Statistics Toolbox Release R2020a, the MathWorks). The base map of the country boundaries was from the Global Administrative Areas dataset (https://gadm.org/download_world.html).

One advantage of our models is that they show more contrasting patterns of EFs than the IPCC default values[13], which differ only for wet climate or intermittently flooded rice (Supplementary Fig. 4 and Supplementary Text 2). Recent N-rate-dependent models[15] produced EF maps with smaller areas of high EF values, yet with a trend of exponentially increasing emissions as N application rate increased that is consistent with our LME models (Supplementary Fig. 5). Management-dependent models[16], considering all variables related to fertilization, irrigation and tillage rather than environmental variables, show the lowest spatial contrasts in EFs, possibly because the effect of N application rates on EFs is offset by other management practices.

Drivers of spatial variation in EFs. We performed multi-model inference[21] to disentangle the relative importance of the drivers that shape EF patterns globally. For all crops, climatic and edaphic variables were the most important drivers, while management-related variables were less important at this scale of analysis (Fig. 2). The link between EFs and environmental variables was consistent with factors known to affect metabolic activities of N-related microbial communities and N_2O production in soils[22]. Large-scale molecular-level investigations reveal that the geographic distributions of nitrifiers and denitrifiers were significantly explained by climatic and edaphic variables[23], controlling N and carbon substrates, oxygen availability, enzymatic activity and metabolic energy sources. Partial correlation analyses between the predicted EFs and drivers using 3.75° moving windows further confirmed such findings at the regional scale (Fig. 3), despite a progressively smaller effect of environmental variables and a larger effect of management-related variables when zooming in the moving windows from 3.75° to 0.75° (Supplementary Fig. 6).

For maize, the spatial variation in EFs was positively associated with soil organic carbon (SOC) in 35% of harvested areas, mainly in the northern latitudes (>45° N, Fig. 3a). This result is in line with the importance of labile carbon substrates for denitrification through mineralization[24]. EFs in an additional 16% and 19% of harvested areas were positively related to humidity index and irrigation fraction, respectively (Fig. 3a). This result is consistent with regional-scale studies[25,26] and manipulation experiments[27] for maize. Higher humidity and more irrigation increase soil moisture (Supplementary Fig. 7), which limits oxygen availability for soil microbes and promotes denitrification[22,26]. This denitrification regime may be dominated by N_2O production, as maize soils rarely reach the soil moisture optimum for N_2O emissions during the growing season[25].

For wheat, the EFs over most (74%) of the harvested area are also correlated with SOC (Fig. 3b), with a few areas where humidity index (3%) and soil pH (11%) are important. Similar results were found for rice. The spatial variability of rice EFs is associated positively with SOC in nearly two-thirds of harvested areas and with clay content in another 21%, mainly in northeastern China and South Asia (Fig. 3c). Increasing clay content in rice fields results in higher porosity and lower water-filled pore space at the same volumetric water content[28], with soil N being more likely to be reduced to N_2O under anaerobic conditions[29]. Other crops are planted across all continents, and their EFs are associated positively with SOC in 35% of harvested areas, mainly in Europe and the tropics (Fig. 3d). We also found a positive correlation between EF and N application rate, mainly in highly fertilized areas such as those growing fruits and vegetables, and between EF and irrigation fraction in water-stressed areas, underscoring the importance of agricultural management practices in shaping N_2O emissions from other crops.

We expected that N application rate would be the most important driver of EF variations as previously emphasized[15], but this was not the case. Field observations that had three or more N levels ($n = 1{,}115$) indicate that the N application rate certainly contributed to higher EFs for major crop and fertilizer types, but it promoted EF values only by 2–7% at under-fertilized sites (for example,

Table 1 | Global direct N_2O emissions and mitigation potentials in cropland (the top ten countries were ranked by mitigation potential)

Country	Emission		Mitigation	
	Quantity (Gg N_2O-N yr^{-1})	Proportion (%)	Quantity (Gg N_2O-N yr^{-1})	Proportion (%)
China	209.3	20.9	95.2	31.3
United States	100.0	10.0	27.7	9.1
India	81.6	8.2	20.8	6.8
Brazil	45.3	4.5	14.8	4.9
Mexico	28.7	2.9	12.1	4.0
Germany	30.6	3.1	9.5	3.1
Pakistan	18.4	1.8	8.4	2.7
Indonesia	32.0	3.2	6.7	2.2
France	31.1	3.1	6.2	2.0
Bangladesh	15.3	1.5	6.2	2.0
Top ten countries	592.3	59.2	207.4	68.2
All countries	1,000.5	100	304.3	100

sub-Saharan Africa) and by 9–24% at over-fertilized sites (for example, North China Plain) (Supplementary Text 3 and Supplementary Fig. 8). This result also suggests that N application rate has a limited contribution in shaping EFs globally. Similar results were found for other agricultural management practices, mainly due to three causes. First, the importance of drivers changes at different spatial scales. Climatic or edaphic variables drive patterns at global and regional scales, but locally, management-related variables tend to be more important. For example, irrigation became the key driver in shaping EFs for maize and other crops when moving windows were shrunk to 0.75° (Supplementary Fig. 6). One or more management practices may thus be the most important drivers of EF within each study site, rather than across them all. Second, across larger scales, environmental variability could influence the farmers' choices of management practices to maintain crop yields[30], which indirectly affects local EFs. Third, according to paired differences analyses[31–33], fertilization, irrigation and tillage practices show diverse effects on EFs, but their combination may offset the effects of each other and thereby weaken their power in shaping EFs. Being able to account for this scaling-up effect with appropriate methods may further alter the perceived importance of agricultural management practices.

Mitigation potential. Combining spatially explicit EFs with N input reduction data allows us to refine N_2O mitigation potential in global cropland. To do so, we developed gridded maps of decreased N input by capping N surplus at a level that maintains crop yield and stays within the proposed planetary boundary. Here, N surplus is defined as N input minus N output in crop yield[9]. This level was defined using data from 361 field trials globally as 48 kg N ha^{-1} for maize, 37 kg N ha^{-1} for wheat, 44 kg N ha^{-1} for rice and 79 kg N ha^{-1} for other crops, after which crop yields tend to plateau (Supplementary Fig. 9a). Additional evidence, which is from a national campaign in China with 20.9 million farmers using integrated soil–crop system management[34], confirmed that crop yield can be increased by ~11% compared with conventional practices even when lowering N surplus to the detected levels (Supplementary Fig. 9b). This scenario decreased the global cropland N surplus from the current 74 to 40 Tg N yr^{-1}, representing a more ambitious boundary than previous estimates[9,35,36] (47–52 Tg N yr^{-1}, Supplementary Fig. 9c). Global mitigation potential from cropland N reduction was then estimated as 0.30 (0.23–1.44) Tg N_2O-N yr^{-1}, which accounted for 30% (17–53%) of global direct emissions of N_2O from cropland and was equivalent to the sum of direct soil emissions from China and the United States combined (Table 1). Non-staple crops contributed 63% of the overall mitigation potential, followed by maize (17%), rice (11%) and wheat (9%), primarily due to lower N use efficiency and higher EF values compared with staple crops.

The N_2O mitigation potentials were unevenly distributed across croplands (Supplementary Fig. 10). We ranked local mitigation potentials from largest to smallest and calculated the cumulative mitigation potential for a given fraction of harvested area (Fig. 4). We found that 20% of the global harvested area for each crop accounts for 50% (47–64%) of the N_2O mitigation potential for maize, 54% (46–64%) for wheat, 43% (36–76%) for rice, 69% (55–82%) for other crops and 65% (49–70%) on average for all crops together. The largest mitigation potentials were found in croplands in the humid subtropical (that is, Cfa and Cwa in the Köppen climate classification), temperate oceanic (Aw) and tropical savannah (wet) climate zones (Cfb), and across gleysol and acrisol soils (Supplementary Fig. 10). The top ten countries together accounted for 68% of the global mitigation potential (Table 1). The countries with the largest mitigation potential were generally those with the highest emissions. Mexico and Pakistan, however, have moderate N_2O emissions but a higher rank for mitigation because of their high EFs and N surplus. Conversely, Canada, Indonesia and France have large emissions but relatively low mitigation potentials because of their relatively modest N surplus.

Interestingly, hotspots of N_2O mitigation potential do not overlap completely with the croplands where N input should be decreased (Supplementary Fig. 11), suggesting that mitigation potential may have a Pareto optimum between high-N-surplus and high-EF regions. For comparison, we quantified mitigation potentials with the same approach but using IPCC Tier 1 defaults[13], N-rate-dependent EFs[15] and management-dependent EFs[16] (Supplementary Text 2) to identify the areas contributing to the same global mitigation potential as our spatially explicit EFs. We find that the best opportunities for mitigation were either overlooked or misjudged broadly by these traditional models (Supplementary Figs. 12–14). Together, these findings suggest that national or regional policy could be improved by using the developed EF maps to better target the local environmental conditions that are critical to assess mitigation potential more accurately and thus drive effective actions.

Limitation of this study and perspective for future work. Our estimates of global cropland N_2O mitigation potential are probably conservative. One reason for this is that the LME models

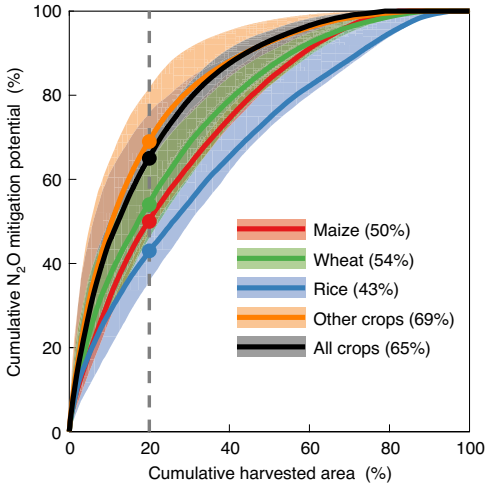

Fig. 4 | Global cropland N_2O mitigation potentials by crop. The shading around each line indicates the 95% CI of mitigation potential due to uncertainties stemming from sampling, modelling and input data. The numbers represent the cumulative proportions of mitigation potential from 20% of the global harvested area by crop (dashed line). The line for all crops represents the cumulative mitigation potential summed for maize, wheat, rice and other crops.

underpredict the measured values of EFs by 13–27% on average (Supplementary Fig. 15), resulting in lower values than previous ones from atmospheric inversion[37] and process-based model ensembles[3]. Furthermore, we focused only on growing-season EFs, as 89% of the field observations compiled in this study did not cover the fallow period. In a sensitivity analysis, we assessed the effect of underrepresented sampling period on the mitigation potential (Supplementary Text 4) and found that neglecting the fallow period had little influence on the pattern of mitigation potentials, but regionally underestimated its magnitude by 30% or more, for wheat in temperate regions (which has a long fallow period when spring thawing occurs[38]) and for rice in the tropics (for which soil moisture declines after harvesting, and N_2O emissions can be stimulated[39]). Finally, we only considered the mitigation potential of direct soil emissions induced by the reduction in N inputs on cropland. In fact, reducing N inputs is also beneficial in mitigating indirect soil emissions from downwind and downstream ecosystems because of less ammonia volatilization and nitrate runoff and leaching.

Our estimates are subject to several sources of uncertainties. The key reason is that microscale variables were less recorded and their effect on local EFs were not fully quantified due to limited understanding of the mechanisms of microbial N_2O production[40]. This is indicated by the result that random effects in the LME models explained more variance in EFs than the fixed effects did (47–74% versus 19–35%, Supplementary Table 3) and contributed most of the estimation uncertainty (Supplementary Fig. 16). Further uncertainty is related to sampling. We acknowledge that the sampling frequency and replication at most sites were too limited to capture the high spatio-temporal variability of N_2O flux, and that the history of control sites is incomplete so that we could not exclude observation data with large legacy fluxes at the control sites. Fortunately, our global compilation of paired control–treatment data used to model EFs may help minimize this bias (Supplementary Text 4). In addition, some global datasets of agricultural management practices have omissions, particularly for fertilization type, timing and placement, enlarging the uncertainty interval of mitigation potentials (Supplementary Text 4).

Additional research that should be prioritized to reduce the uncertainty of mitigation potential is listed as follows. First, high-frequency, year-round observations of direct and indirect N_2O emissions would substantially improve our approach. Second, more detailed records on the status of control sites are needed for filtering observation data to avoid bias in the quantification of EFs. This additional information should include the year when each control site was first fertilized before the experiment and its level of soil residual N. Third, resolving and measuring microscale biophysical characteristics is a promising route to quantify their roles in explaining cross-site variation in N_2O emissions. In addition to the collection of subnational statistics, more efforts should be made to improve the accuracy of management datasets by taking advantage of high-resolution satellite data and large-scale machine learning (for example, see ref. [41]).

Overall, our global, five-minute-resolution maps of N_2O EFs and analysis of mitigation potentials provide information for guiding field observation network design, refining national emission inventories and further targeting actions towards mitigation priority areas. Achieving global N_2O mitigation in cropland will be challenging for a variety of reasons, including the difficulty of implementing precise agricultural management technologies across millions of hectares and verifying their effectiveness[42]. Measures in addition to N input reduction are therefore needed to increase mitigation potentials. Possible options include the application of so-called enhanced fertilizers (for example, nitrification inhibitor) to lower EFs[43], the introduction of high-yield cultivars to improve N use efficiency[44] and spatial reallocation of N resources adapted to local conditions that favour crop production while inhibiting N_2O emissions[45]. It is also critical to optimize measures designed to mitigate N_2O emissions while avoiding other N losses (for example, NH_3 volatilization[46], N runoff and leaching[47]) and to achieve a complete closure of the N budgets of agricultural soils[48]. The success of N_2O mitigation depends not only on strong multi-national collaborations, investment incentives, schemes for providing mitigation revenues to farmers, but also on overcoming technical and socio-economic barriers[49].

Methods

Observation dataset. We compiled a global observation dataset of cropland N_2O emissions from the literature and online data repositories (Supplementary Table 4), including 4,924 emission flux and 2,698 EF records from chamber-based field experiments for global croplands. We then excluded records that (1) lacked replicates or zero-N control, (2) used enhanced fertilizers either treated with inhibitors or coated, or (3) were sampled with a frequency of less than one flux measurement per week or for a duration of less than one crop growing season. We also excluded records for which information on management practices was neither derived from published statements nor supplemented by contacting the authors. This yielded a dataset of 1,507 EF values from 249 experiments that span 31 countries and 234 sites (Supplementary Text 1). The full dataset can be split into four subsets by crop—that is, 458 measurements for maize, 313 for wheat, 284 for rice and 452 for the other crops.

For each record, six categories of information were collected. First, N_2O emissions included the observed direct N_2O flux under different N levels. The EF (the percentage of N applied) was calculated for a non-zero N application rate (N_{ij}) as $EF_{ij} = (E_{ij} - E_{0j})/N_{ij}$, where i is the index of N-input levels, j is the index of crop type, E_{ij} (kg N_2O-N ha^{-1}) is the direct N_2O flux due to the application of N inputs (that is, synthetic fertilizer, manure, crop residues or their combinations) and other unquantified sources (for example, atmospheric deposition or soil residual N), and E_0 is the direct N_2O flux at a zero-N control site due to other unquantified sources. Second, crop types considered in the database included wheat, maize, rice and other crops. Third, climatic variables included mean daily air temperature, cumulative precipitation and humidity index (that is, the ratio of precipitation to evapotranspiration) within the measurement period. Fourth, soil variables included bulk density, soil pH, soil clay content and SOC content. Fifth, management-related variables included fertilization (that is, N application rate, type, frequency and placement), irrigation fraction (that is, rainfed or irrigated for upland crops, and continuously or intermittently flooded for rice) and tillage fraction (that is, no-till or till). Sixth, experimental parameters included location (that is, latitude and longitude), sampling frequency (that is, times per week), duration (that is, days from starting to ending dates) and replications (that is, the

standard deviation of EF derived from replicates). The definition and unit of each variable can be found in Supplementary Table 5.

For each of the four crop types, we used the method of van den Hoogen et al.[50] to investigate how well our newly compiled dataset spread throughout the full multivariate covariate space (all environmental and management-related variables) of the global layers. Interpolation percentage is defined by estimating how adequately our dataset captured the multivariate covariate space of the global layers. Additional comparisons indicate that the interpolation capability of our global dataset was higher than that in previous studies[15,19,20], regardless of crop group (Supplementary Fig. 2).

LME modelling. We prefer the LME model to the random forest (RF) for predicting EFs for four main reasons. First, we aimed to generate an interpretable model of EFs in response to environmental and management-related variables. This would enable data-deficient countries to move away from the IPCC Tier 1 default values towards refining their national greenhouse gas inventories. Second, we needed to investigate the overall uncertainties around the estimates and predictions of EFs and mitigation potentials. Third, the performance of the LME model evaluated with block-cross-validation of EFs was comparable to that of the RF, regardless of crop type (Supplementary Table 6). Last, to confirm that our predictions were robust against the model algorithm used, the RF model was also used to estimate global cropland N_2O EFs and mitigation potentials. The results indicate no perceivable differences in magnitude or pattern between the predictions from LME and RF models (Supplementary Figs. 17 and 18). The details of the model setup and validation are provided below.

We used the observation data to determine whether it was necessary to model EFs by crop type. A non-parametric Wilcoxon test indicated significant differences among crops ($P < 0.05$, Supplementary Fig. 19). Including crop type in the LME model resulted in better model performance than excluding it (Supplementary Table 7). However, a single LME model cannot quantify the inter-crop differences in the sensitivity of EFs to predictors revealed by partial correlation analysis (Supplementary Fig. 20). We therefore constructed LME models separately for maize, wheat and rice as the staple crops globally, while grouping other crops together due to the lack of sufficiently available observations for each of them.

For each crop type, we checked whether one or two random-effects terms (such as site identity) were required. An analysis of variance test indicated that models with random effects outperformed those with fixed effects only (Supplementary Table 8). Including random-effects terms also made the coverage probability of prediction interval (90–95%) very close to the theoretical value of 95% (Supplementary Table 9), suggesting that random-effects terms are important for quantifying prediction intervals of EFs and mitigation potential. Although models including two random effects in the intercept and slope outperformed those including one random effect in the intercept only for rice and other crops (Supplementary Table 9), it is also meaningful to include two random effects for each crop to account for random variability of EFs as previously recognized[14].

We screened continuous variables to avoid collinearity between them. Prior to variable selection, the outliers of EFs were filtered on the basis of the *Z*-score outlier test (4 measurements for maize and 27 measurements for other crops). We took the natural logarithmic transformation of EF due to its skewness to fit the LME model, and we centred and scaled the continuous variables (that is, climatic and edaphic variables and N application rate) to have a mean of zero and a variance of 1.0. We considered three forms (linear, logarithmic and exponential) of continuous variables to allow for a nonlinear effect of predictors on EFs, as is commonly found in field studies[15,26]. We also considered two-way interactions within soil and climate themes as well as between soil, climate and N application rate. For each iteration of variable selection, observation uncertainty due to inter-site differences in sampling frequency and replication was used to weight the LME models. The number of variables was reduced until all remaining variables had a variance inflation factor <10. To avoid over-fitting, backward variable selection using the Akaike information criterion (AIC) was implemented with the R package lme4[51] (v.1.1-21). Eventually, the LME model for each crop retained most of the continuous variables and all categorical variables as fixed-effects terms, including bulk density; clay content; SOC; soil pH; precipitation; temperature; humidity index; fertilizer application rate, type, frequency and placement; irrigation fraction; tillage fraction; and sampling duration, but interactions were removed (Supplementary Table 5). The model also included site identity in the intercept and the slope of N rate as random-effects terms.

Finally, additional analyses were conducted to evaluate the robustness and predictability of each of the four LME models. First, leave-one-out cross-validation was performed by separating training and testing sets on the basis of site blocks. The bulk of the data were trained, and a holdout site was used to test the model. The root mean squared error (RMSE) was calculated for all predicted data and for terciles of the observations after the leave-one-out cross-validation. Second, RF models were fitted by including all the same predictors as in the LME models to assess the robustness of our findings. Last, the influences of filling the edaphic and climatic variables from CRU TS v.4.03 (https://doi.org/10.5285/10d3e3640f004c578403419aac167d82) and HWSD v.1.2 (http://dare.iiasa.ac.at/44/2/HWSD.zip) were tested. LME models were produced again using edaphic and climatic variables totally extracted from global datasets. The same variable selection procedure was applied as described above, and the alternative model was identified for each of the four crops. The results from leave-one-out cross-validation suggest that model performance was insensitive to the data source for climatic or edaphic variables and robust against the algorithm used (LME versus RF), regardless of crop type (Supplementary Table 6).

Global predictions. The global patterns of crop-specific N_2O EFs in 2000 were predicted using the LME models by crop at five-arcminute spatial resolution. The input data included the global gridded dataset of climate, soil, fertilization, irrigation and tillage for 26 crop types (Supplementary Table 10). Note that the LME model for other crops was used for all crop types except the three staple crops. Climate data over the growing season for a given crop type were acquired from CRU TS v.4.03 (0.5° resolution), where the growing season in each grid cell was identified as the period between the planting and harvesting dates obtained from Sacks et al.[52]. Soil data were acquired directly from the HWSD v.1.2 (30-arcsecond resolution). Both climate and soil properties were re-gridded at a resolution of 5′ × 5′, while the remaining datasets were specifically developed for this study.

Fertilization. We provided a global gridded, crop-specific fertilization dataset including the rate, type, frequency and placement of N inputs in 2000. First, we collected synthetic fertilizers from 15,790 global administrative units (that is, 14,855 counties, 740 provinces or states, and 195 countries). We downloaded the five-arcminute gridded data on manure applied to croplands from Zhang et al.[53] and the national data on crop residues from the Food and Agriculture Organization (FAO)[54] and then resampled them into 15,790 administrative units. We then allocated these N inputs into 26 crops for each administrative unit, on the basis of the previously developed proportions of fertilizer uses by crop from EarthStat[45]. Synthetic fertilizers were further disaggregated into four products (urea, ammonium nitrate, calcium ammonium nitrate and other fertilizers) on the basis of the provincial or state-level statistics for the United States[55], China[56] and India[57] and IFASTAT's country-level statistics[58] for the other countries. N application rates by crop and fertilizer were then disaggregated into grid maps at five-arcminute spatial resolution following the global harvested area distributions[59] within each of the administrative units. The maximum N application rate was set at 1,000 kg N ha^{-1}. We next determined the frequency (one or multiple times) of fertilizer applications on the basis of field surveys and previous literature (Supplementary Table 11). Finally, we quantified the fraction of different placement methods in each grid cell depending on fertilizer type, tillage practice, crop type and growing period[46]. Manure and crop residues that are usually applied before sowing or transplanting and afterwards were incorporated linearly in response to tillage fraction (or 100% for non-tillage). Anhydrous ammonia and N solutions, which are commonly injected, were placed deep or totally incorporated into soils. The rest of the synthetic fertilizers were usually incorporated for vegetables and fruits due to their higher economic return and planting density, but were applied by broadcasting techniques during topdressing for other crops if applicable.

Tillage. We downloaded the global gridded crop-specific non-tillage dataset (5′ × 5′) from Porwollik et al[60]. We then compared the crop-specific non-tillage area with the publicly accessible data from Canada, Brazil, Australia, China and the United States (Supplementary Fig. 21), representing the top and low adopters of non-tillage practices in the world. This dataset reproduced well the crop-specific proportion of non-tillage area at the national scale (slope = 0.76, RMSE = 0.06), but at the subnational scale, it tended to concentrate non-tillage area in a few regions instead of a more homogenous spread as reported (slope = 0.54, RMSE = 0.19).

Irrigation. The MIRCA2000 dataset[61] provides both irrigated and rainfed areas by crop at five-arcminute spatial resolution. This dataset maximizes consistency with the cropland data and draws from FAO's AQUASTAT database, though it may underestimate irrigated areas by 18% compared with the satellite-based approach[62]. We calculated the fractions of irrigated cropland area separately for 26 crop types. We further divided irrigated rice into continuously and intermittently flooded systems. For the top 12 rice production countries except China, country-scale data were acquired from ALGAS reports[63] and local statistics. For China, the provincial data were compiled from Zhang et al.[64], according to the proportion of year-round waterlogged rice paddies in hilly regions. For the rest of the world, the data were determined by Carlson et al.[65]. Rainfed rice can be classified into shallow (<30 cm water depth), intermediate (30–100 cm) and deepwater (>100 cm). The shallow type is drought-prone or both drought- and submergence-prone, and generally it is equivalent to the intermittently flooded system, while the intermediate and deepwater types are equivalent to the continuously flooded system[65].

Relative importance. We performed multi-model inference[21] based on the AIC to estimate which predictor variables were the most important in driving patterns of cropland N_2O EFs at the global scale. This method surpasses the single best AIC model, which may miss important variables, and accounts for uncertainty not only in the parameter estimates but also in the model. We first generated a full submodel set from the global model using the dredge function implemented in

the R package MuMIn[66] based on the LME model. We obtained a set of top models using a cut-off of cumulative weight (cumsum(weight) ≤ 0.9999) and used the model.avg function to estimate model parameters on the basis of the top models. We then calculated the relative importance of individual variables by summing the Akaike weights across all models that contained the candidate variable. It should be noted that this analysis was for growing-season EFs. To assess the robustness of our results, a similar analysis was done for whole-year EFs on the basis of a limited subset of the observation dataset ($n = 168$, Supplementary Text 4).

In addition, we performed a partial correlation analysis[67] between EFs and variables to identify the dominant driver of variation in EFs at the regional scale. This analysis was carried out with 3.75°-by-3.75° moving windows. The resolution of the data was 5′ by 5′—that is, the surrounding 2,025 pixels were used for each 5′ pixel. We first calculated the coefficient and significance of partial correlation in each pixel, and we then identified the dominant driver with the largest absolute value of the correlation coefficient. To assess the robustness of this finding, a similar analysis was done with moving windows at higher spatial resolutions—that is, 0.75° by 0.75°, 1.75° by 1.75° and 2.75° by 2.75°.

Mitigation potential assessments. We estimated the global crop-specific N_2O mitigation potential from the decrease in N application rate (N_{rate}) by capping N surplus to the limit (N_L) without compromising crop yield or breaching the bounds for acceptable air and water quality. To do so, we accessed a publicly available CGIAR dashboard[68] to extract the global observations of N yield and N surplus along with N_{rate} gradients. This yielded a dataset of 144 records for maize, 34 records for wheat, 57 records for rice and 126 records for other crops. We then used linear-to-plateau models to detect the value of N_L by crop (mean and 95% CI) above which crop yields show stagnation (significance level, $P < 0.05$). The models were developed by implementation in the R package easyreg[69]. To test the feasibility of achieving the detected levels of N surplus while not compromising crop yield at the regional scale, we provided evidence from China's national campaign in 2005–2015 that encouraged 20.9 million farmers to adopt integrated soil–crop system management technologies for greater crop yield and reduced environmental pollution[34]. We finally calculated the reduced N surplus summed from global cropland to test whether it stays within the planetary boundary as previously estimated[9,35,36].

Note that our objective is to assess the mitigation potential for direct N_2O emissions, not to optimize mitigation measures considering technical barriers and marginal abatement costs. First, we estimated the global cropland N surplus (N_{sur}) by grid cell and crop type as the sum of N inputs (fertilizers, manure, biological fixation and atmospheric depositions) minus N outputs (N_{yield}). Gridded data on biologically fixed N and N_{yield} were calculated on the basis of the crop yield data from the EarthStat databases (5′ × 5′) and the crop-specific parameters (N fixation rate and N content by crop) from a previous study[9]. Atmospheric N deposition rates were extracted from the IGAC/SPARC Chemistry-Climate Model Initiative (0.5° × 0.5°) over global croplands[3]. Second, we estimated the decreased N application rate everywhere if applicable by subtracting the difference between N_{sur} and the detected level—that is, $N_{rate} - \Delta N_{sur}$, where $\Delta N_{sur} = N_{sur} - N_L$ (Supplementary Fig. 22). Finally, we used the four LME models to re-estimate crop-specific EFs using the decreased N application rate and thereby estimated the direct cropland N_2O emissions and mitigation potentials by crop type.

Uncertainty estimation. A Monte Carlo simulation was used to estimate the overall uncertainty for predicting EFs and mitigation potentials. To generate a proper prediction interval, our estimates accounted for three sources of uncertainty: the fixed coefficients, the random coefficients and the input data. The uncertainty of EFs taken from sampling frequency and replication was reflected by the first source of variation, while the uncertainty of EFs from unquantified sources was reflected by the other two sources. Each crop-specific LME model was run by randomly generating the fixed and random coefficients from their fitted multivariate normal distributions and the climate and soil variables, N application rate, fraction of fertilizer types, fraction of fertilizer placement, tillage fraction and irrigated fraction following independent normal distributions with the same mean as our dataset and a standard deviation of the absolute difference between the dataset used in this study and other global datasets (Supplementary Table 12). Fertilizer frequency (one or more times) was also randomly selected following the Bernoulli (two-point) distribution. We then calculated the predicted values from the LME models through 1,000 iterations so that the mean and 2.5% and 97.5% quantiles could be constructed with the 95% prediction interval. We also broke down the uncertainty of EFs and mitigation potential per source of uncertainty, suggesting that the uncertainty of estimation stemmed mainly from the random coefficients (Supplementary Fig. 16).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The global cropland N_2O emission observation datasets compiled for this study are available in Supplementary Data 1. The global input datasets of fertilization, irrigation and tillage practices developed for this study are available at https://doi.org/10.6084/m9.figshare.14842965. The model outputs, including global gridded maps of N_2O EFs and mitigation potential (including means and 95% CIs), are available at https://doi.org/10.6084/m9.figshare.14844069. The climate data are available at https://crudata.uea.ac.uk/cru/data/hrg/. The soil data are available at https://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/. The harvested area and crop yield data are available at http://www.earthstat.org/harvested-area-yield-175-crops/. Source data are provided with this paper.

Code availability

The computer code for statistics, global prediction and uncertainty estimation is available at https://doi.org/10.6084/m9.figshare.16353480.

Received: 30 January 2021; Accepted: 9 September 2021;
Published online: 21 October 2021

References

1. Quan, Z., Zhang, X., Fang, Y. & Davidson, E. A. Different quantification approaches for nitrogen use efficiency lead to divergent estimates with varying advantages. *Nat. Food* **2**, 241–245 (2021).
2. *Emissions from Agriculture and Forest Land: Global, Regional and Country Trends 1990–2019* FAOSTAT Analytical Brief Series No. 25 (FAO, 2021); http://www.fao.org/3/cb5293en/cb5293en.pdf
3. Tian, H. et al. A comprehensive quantification of global nitrous oxide sources and sinks. *Nature* **586**, 248–256 (2020).
4. Clark, M. A. et al. Global food system emissions could preclude achieving the 1.5° and 2 °C climate change targets. *Science* **370**, 705–708 (2020).
5. Portmann, R. W., Daniel, J. S. & Ravishankara, A. R. Stratospheric ozone depletion due to nitrous oxide: influences of other gases. *Phil. Trans. R. Soc. B* **367**, 1256–1264 (2012).
6. *The Future of Food and Agriculture: Alternative Pathways to 2050* (FAO, 2018); http://www.fao.org/3/I8429EN/i8429en.pdf
7. Bodirsky, B. L. et al. Reactive nitrogen requirements to feed the world in 2050 and potential to mitigate nitrogen pollution. *Nat. Commun.* **5**, 3858 (2014).
8. de Vries, W., Kros, J., Kroeze, C. & Seitzinger, S. P. Assessing planetary and regional nitrogen boundaries related to food security and adverse environmental impacts. *Curr. Opin. Environ. Sustain.* **5**, 392–402 (2013).
9. Zhang, X. et al. Managing nitrogen for sustainable development. *Nature* **528**, 51–59 (2015).
10. Decock, C. Mitigating nitrous oxide emissions from corn cropping systems in the midwestern US: potential and data gaps. *Environ. Sci. Technol.* **48**, 4247–4256 (2014).
11. Groffman, P. M. et al. Challenges to incorporating spatially and temporally explicit phenomena (hotspots and hot moments) in denitrification models. *Biogeochemistry* **93**, 49–77 (2009).
12. Ogle, S. M., Butterbach-Bahl, K., Cardenas, L., Skiba, U. & Scheer, C. From research to policy: optimizing the design of a national monitoring system to mitigate soil nitrous oxide emissions. *Curr. Opin. Environ. Sustain.* **47**, 28–36 (2020).
13. Calvo Buendia, E. et al. (eds) *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories* (IPCC, 2019).
14. Gerber, J. S. et al. Spatially explicit estimates of N_2O emissions from croplands suggest climate mitigation opportunities from improved fertilizer management. *Glob. Change Biol.* **22**, 3383–3394 (2016).
15. Shcherbak, I., Millar, N. & Robertson, G. P. Global metaanalysis of the nonlinear response of soil nitrous oxide (N_2O) emissions to fertilizer nitrogen. *Proc. Natl Acad. Sci. USA* **111**, 9199–9204 (2014).
16. Wang, Q. et al. Data-driven estimates of global nitrous oxide emissions from croplands. *Natl Sci. Rev.* **7**, 441–452 (2020).
17. Dorich, C. D. et al. Improving N_2O emission estimates with the global N_2O database. *Curr. Opin. Environ. Sustain.* **47**, 13–20 (2020).
18. Heffer, P., Gruère, A. & Roberts, T. *Assessment of Fertilizer Use by Crop at the Global Level* (International Fertilizer Association, 2014).
19. Stehfest, E. & Bouwman, L. N_2O and NO emission from agricultural fields and soils under natural vegetation: summarizing available measurement data and modeling of global annual emissions. *Nutr. Cycl. Agroecosyst.* **74**, 207–228 (2006).
20. Wang, Y. et al. Soil pH as the chief modifier for regional nitrous oxide emissions: new evidence and implications for global estimates and mitigation. *Glob. Change Biol.* **24**, E617–E626 (2018).
21. Grueber, C. E., Nakagawa, S., Laws, R. J. & Jamieson, I. G. Multimodel inference in ecology and evolution: challenges and solutions. *J. Evol. Biol.* **24**, 699–711 (2011).
22. Butterbach-Bahl, K., Baggs, E. M., Dannenmann, M., Kiese, R. & Zechmeister-Boltenstern, S. Nitrous oxide emissions from soils: how well do we understand the processes and their controls? *Phil. Trans. R. Soc. B* **368**, 20130122 (2013).
23. Nelson, M. B., Martiny, A. C. & Martiny, J. B. H. Global biogeography of microbial nitrogen-cycling traits in soil. *Proc. Natl Acad. Sci. USA* **113**, 8033–8040 (2016).

24. Venterea, R. T. Nitrite-driven nitrous oxide production under aerobic soil conditions: kinetics and biochemical controls. *Glob. Change Biol.* **13**, 1798–1809 (2007).
25. Griffis, T. J. et al. Nitrous oxide emissions are enhanced in a warmer and wetter world. *Proc. Natl Acad. Sci. USA* **114**, 12081–12085 (2017).
26. Parn, J. et al. Nitrogen-rich organic soils under warm well-drained conditions are global nitrous oxide emission hotspots. *Nat. Commun.* **9**, 1135 (2018).
27. Wu, G. et al. Effects of soil warming and increased precipitation on greenhouse gas fluxes in spring maize seasons in the North China Plain. *Sci. Total Environ.* **734**, 139269 (2020).
28. Balaine, N. et al. Changes in relative gas diffusivity explain soil nitrous oxide flux dynamics. *Soil Sci. Soc. Am. J.* **77**, 1496–1505 (2013).
29. Hall, S. J., Reyes, L., Huang, W. & Homyak, P. M. Wet spots as hotspots: moisture responses of nitric and nitrous oxide emissions from poorly drained agricultural soils. *J. Geophys. Res. Biogeosci.* **123**, 3589–3602 (2018).
30. Paustian, K. et al. Climate-smart soils. *Nature* **532**, 49–57 (2016).
31. Jiang, Y. et al. Water management to mitigate the global warming potential of rice systems: a global meta-analysis. *Field Crops Res.* **234**, 47–54 (2019).
32. Mei, K. et al. Stimulation of N_2O emission by conservation tillage management in agricultural lands: a meta-analysis. *Soil Tillage Res.* **182**, 86–93 (2018).
33. Xia, L. et al. Can knowledge-based N management produce more staple grain with lower greenhouse gas emission and reactive nitrogen pollution? A meta-analysis. *Glob. Change Biol.* **23**, 1917–1925 (2017).
34. Cui, Z. et al. Pursuing sustainable productivity with millions of smallholder farmers. *Nature* **555**, 363–366 (2018).
35. EU Nitrogen Expert Panel *Nitrogen Use Efficiency (NUE): An Indicator for the Utilization of Nitrogen in Agriculture and Food Systems* (Wageningen University, Alterra, 2015).
36. Springmann, M. et al. Options for keeping the food system within environmental limits. *Nature* **562**, 519–525 (2018).
37. Thompson, R. L. et al. Acceleration of global N_2O emissions seen from two decades of atmospheric inversion. *Nat. Clim. Change* **9**, 993–998 (2019).
38. Wagner-Riddle, C. et al. Globally important nitrous oxide emissions from croplands induced by freeze–thaw cycles. *Nat. Geosci.* **10**, 279–283 (2017).
39. Kritee, K. et al. High nitrous oxide fluxes from rice indicate the need to manage water for both long- and short-term climate impacts. *Proc. Natl Acad. Sci. USA* **115**, 9720–9725 (2018).
40. Kravchenko, A. N. et al. Hotspots of soil N_2O emission enhanced through water absorption by plant residue. *Nat. Geosci.* **10**, 496–500 (2017).
41. Phalke, A. R. et al. Mapping croplands of Europe, Middle East, Russia, and Central Asia using Landsat, Random Forest, and Google Earth Engine. *ISPRS J. Photogramm. Remote Sens.* **167**, 104–122 (2020).
42. Roe, S. et al. Contribution of the land sector to a 1.5 °C world. *Nat. Clim. Change* **9**, 817–828 (2019).
43. Lam, S. K., Suter, H., Mosier, A. R. & Chen, D. Using nitrification inhibitors to mitigate agricultural N_2O emission: a double-edged sword? *Glob. Change Biol.* **23**, 485–489 (2017).
44. Liu, Y. et al. Genomic basis of geographical adaptation to soil nitrogen in rice. *Nature* **590**, 600–605 (2021).
45. Mueller, N. D. et al. Closing yield gaps through nutrient and water management. *Nature* **490**, 254–257 (2012).
46. Zhan, X. et al. Improved estimates of ammonia emissions from global croplands. *Environ. Sci. Technol.* **55**, 1329–1338 (2021).
47. Jiang, W. et al. Is rice field a nitrogen source or sink for the environment? *Environ. Pollut.* **283**, 117122 (2021).
48. Zhang, X. et al. Quantification of global and national nitrogen budgets for crop production. *Nat. Food* **2**, 529–540 (2021).
49. Kanter, D. R. et al. Nitrogen pollution policy beyond the farm. *Nat. Food* **1**, 27–32 (2020).
50. van den Hoogen, J. et al. Soil nematode abundance and functional group composition at a global scale. *Nature* **572**, 194–198 (2019).
51. Bates, D., Maechler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
52. Sacks, W. J., Deryng, D., Foley, J. A. & Ramankutty, N. Crop planting dates: an analysis of global patterns. *Glob. Ecol. Biogeogr.* **19**, 607–620 (2010).
53. Zhang, B. et al. 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 (2017).
54. *FAOSTAT Data* (FAO, accessed 18 June 2020); http://www.fao.org/faostat/en/#data/GT (*Emissions—Agriculture*); http://www.fao.org/faostat/en/#data/RL (*Land Use*)
55. Cao, P., Lu, C. & Yu, Z. Historical nitrogen fertilizer use in agricultural ecosystems of the contiguous United States during 1850–2015: application rate, timing, and fertilizer types. *Earth Syst. Sci. Data* **10**, 969–984 (2018).
56. *Compilation Data of National Agricultural Product Cost and Benefit for China (the Ratio of Each Synthetic Fertilizer Product Applied for Each Crop in 2001 and 2005)* (China Statistics Press, 2001 and 2005); http://data.cnki.net/Yearbook/Single/N2010090051
57. Aneja, V. P. et al. Reactive nitrogen emissions from crop and livestock farming in India. *Atmos. Environ.* **47**, 92–103 (2012).
58. *IFASTAT Data* (International Fertilizer Association, accessed 18 June 2020); https://www.ifastat.org/databases/plant-nutrition
59. Monfreda, C., Ramankutty, N. & Foley, J. A. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Glob. Biogeochem. Cycles* **22**, GB1022 (2008).
60. Porwollik, V., Rolinski, S., Heinke, J. & Müller, C. Generating a rule-based global gridded tillage dataset. *Earth Syst. Sci. Data* **11**, 823–843 (2019).
61. Portmann, F. T. *Global Estimation of Monthly Irrigated and Rainfed Crop Areas on a 5 Arc-Minute Grid*. PhD thesis, Univ. Frankfurt (2011).
62. Meier, J., Zabel, F. & Mauser, W. A global approach to estimate irrigated areas—a comparison between different data and statistics. *HESS* **22**, 1119–1133 (2018).
63. *Asia Least-Cost Greenhouse Gas Abatement Strategy (ALGAS)* (Asian Development Bank, Global Environment Facility and United Nations Development Program, 1998).
64. Zhang, W., Yu, Y., Huang, Y., Li, T. & Wang, P. Modeling methane emissions from irrigated rice cultivation in China from 1960 to 2050. *Glob. Change Biol.* **17**, 3511–3523 (2011).
65. Carlson, K. M. et al. Greenhouse gas emissions intensity of global croplands. *Nat. Clim. Change* **7**, 63–68 (2017).
66. Barton, K. Mu-MIn: Multi-Model Inference. R package version 0.12.2/r18 (2009); http://R-Forge.R-project.org/projects/mumin/
67. Beer, C. et al. Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science* **329**, 834–838 (2010).
68. Maaz, T. M. et al. Meta-analysis of yield and nitrous oxide outcomes for nitrogen management in agriculture. *Glob. Change Biol.* **27**, 2343–2360 (2021).
69. Arnhold, E. R-environment package for regression analysis. *Pesqui. Agropecu. Bras.* **53**, 870–873 (2018).

Acknowledgements

This study was supported by the National Natural Science Foundation of China (grant nos 41671464 to F.Z. and 41830751 to X.J.), the China Postdoctoral Science Foundation (grant no. 2019M660301 to X.C.), the US National Science Foundation (grant no. 1903722 to H.T.) and the French ANR under the CLAND 'Investissements d'avenir' programme (grant no. ANR-16-CONV-0003 to P.C.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. We thank Z. Shang from the University of Aberdeen for sharing the dataset that includes growing-season and whole-year EFs and the fallow EF models. We acknowledge field experimentalists for publishing or sharing chamber-based observations of N_2O flux. We acknowledge FAO and IFA for national statistics on synthetic fertilizers, and we thank 38 agencies for subnational statistics on synthetic fertilizers. The views expressed in this publication are those of the authors and do not necessarily reflect the views or policies of FAO. We also acknowledge IIASA, CGIAR, the University of East Anglia, the University of Minnesota and V. Porwollik from the Potsdam Institute for Climate Impact Research for sharing other model input data.

Author contributions

F.Z. designed the study. X.C., X.N. and F.Z. performed all computational analyses. F.Z. and X.C. drafted the paper. X.C., Q.W., W.A., X. Zhan and Y.B. collected the data and prepared the figures and tables. All authors reviewed and commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43016-021-00384-9.

Correspondence and requests for materials should be addressed to Feng Zhou.

Peer review information *Nature Food* thanks David Makowski, Klaus Butterbach-Bahl and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): FENG ZHOU

Last updated by author(s): Aug 20, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Microsoft Excel Office 2019
Data analysis	R version 3.6.0 AND MATLAB R2016b

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Global cropland-N2O emission observation datasets compiled for this study are available online from Supplementary Data 1. Global input datasets of fertilization, irrigation, and tillage practices developed for this study are available from https://doi.org/10.6084/m9.figshare.14842965. Model outputs, including global gridded maps of N2O EF and mitigation potential (including means and 95% CI) are available from https://doi.org/10.6084/m9.figshare.14844069. Climate data is available from https://crudata.uea.ac.uk/cru/data/hrg/. Soil data is available from https://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/. Harvested area and crop yield data are available from http://www.earthstat.org/harvested-area-yield-175-crops/. All display items that support the findings of this study are available in the Source Data.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	*Describe how sample size was determined, detailing any statistical methods used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data exclusions	*Describe any data exclusions. If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Replication	*Describe the measures taken to verify the reproducibility of the experimental findings. If all attempts at replication were successful, confirm this OR if there are any findings that were not replicated or cannot be reproduced, note this and describe why.*
Randomization	*Describe how samples/organisms/participants were allocated into experimental groups. If allocation was not random, describe how covariates were controlled OR if this is not relevant to your study, explain why.*
Blinding	*Describe whether the investigators were blinded to group allocation during data collection and/or analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*
Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This study aims to address three key questions: (i) what is the heterogeneity of EF across global cropland? (ii) how strongly do these variables influence variations in EF at global and regional scales? and (iii) to what extent can direct soil emissions of N2O be mitigated while maintaining crop yield?

Research sample	This study provides a dataset of 1,507 EF values from 249 experiments that span 31 countries and 234 sites. Climate data over the growing season were acquired from the CRU TS v4.0354 (0.5-degree resolution), where the growing season in each grid cell was identified as the period between the planting and harvesting dates obtained from Sacks et al.. Soil data were acquired directly from the HWSD v1.255 (30-arc-second resolution). Both climate and soil properties were re-gridded at a resolution of 5 5 , while fertilization, irrigation, and tillage datasets were specifically developed for this study.
Sampling strategy	Multiple analyses were conducted to evaluate the robustness and predictability of each of four LME models using the global complication of EF data. First, leave-one-out cross validation was performed by separating training and testing sets based on site blocks. The bulk of the data were trained and a holdout site was used to test the model. The root mean squared error (RMSE) was calculated for all predicted data and for terciles of the observations following the leave-one-out cross-validation. Second, random forest (RF) models were fitted by including all the same predictors as in the LMEs to assess the robustness of our findings. Last, the influences of filling edaphic and climatic variables from CRU TS v4.0354 and HWSD v1.255 were tested. LME models were produced again using edaphic and climatic variables totally extracted from global datasets54,55. The same variable selection procedure was applied as described above, and the alternative model was identified for each of the four crops.
Data collection	We compiled a global observation dataset of cropland-N2O emissions from the literature and online data repositories (Supplementary Table 3), including 4,924 emission flux and 2,698 EF records from chamber-based field experiments for global croplands. We then excluded the records which (i) lacked replicates or zero-N control, (ii) used enhanced fertilizers either treated with inhibitors or coated, (iii) were sampled with the frequency of less than one flux measurement per week or the duration of less than a crop growing season. We also excluded the records for which information of management practices were neither derived from published statements nor supplemented by contacting the authors. This yielded a dataset of 1,507 EF values from 249 experiments that span 31 countries and 234 sites.
Timing and spatial scale	Chamber-based field observations of EF were conducted in 234 sites and 31 countries from 1980 to 2018. The global patterns of crop-specific N2O EFs and mitigation potentials were predicted using four LME models in 2000 at 5-arc-minute spatial resolution.
Data exclusions	Prior to variable selection, the outliers of EFs were filtered based on the Z-Score outlier test52 (4 measurements for maize, and 27 measurements for other crops).
Reproducibility	Our study is an integrated study mainly based on model simulation ,statistical data and field observation data. Our results can be reproduced when following the described methods and data.
Randomization	This is not relevant to our study because our work is not an "experimental" study but an integrated data analysis. We used field observation data and publicly available statistical data to do the integrated analysis.
Blinding	Blinding is not possible in our study. Because we have no choice to collect the subnational statistical data or field observation data (just according to key words) considering the blinding principle

Did the study involve field work? ☐ Yes ☒ No

Field work, collection and transport

Field conditions	*Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*
Location	*State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*
Access & import/export	*Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*
Disturbance	*Describe any disturbance caused by the study and how it was minimized.*

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used — *Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.*

Validation — *Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.*

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s) — *State the source of each cell line used.*

Authentication — *Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.*

Mycoplasma contamination — *Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.*

Commonly misidentified lines (See ICLAC register) — *Name any commonly misidentified cell lines used in the study and provide a rationale for their use.*

Palaeontology and Archaeology

Specimen provenance — *Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).*

Specimen deposition — *Indicate where the specimens have been deposited to permit free access by other researchers.*

Dating methods — *If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight — *Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals — *For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.*

Wild animals — *Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*

Field-collected samples — *For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*

Ethics oversight — *Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics — *Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*

Recruitment — *Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*

Ethics oversight — *Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☐	☐	Public health
☐	☐	National security
☐	☐	Crops and/or livestock
☐	☐	Ecosystems
☐	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☐	☐	Demonstrate how to render a vaccine ineffective
☐	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☐	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☐	☐	Increase transmissibility of a pathogen
☐	☐	Alter the host range of a pathogen
☐	☐	Enable evasion of diagnostic/detection modalities
☐	☐	Enable the weaponization of a biological agent or toxin
☐	☐	Any other potentially harmful combination of experiments and agents

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	*For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*
Files in database submission	*Provide a list of all files available in the database submission.*
Genome browser session (e.g. UCSC)	*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates	*Describe the experimental replicates, specifying number, type and replicate agreement.*
Sequencing depth	*Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and*

Sequencing depth	*whether they were paired- or single-end.*
Antibodies	*Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Peak calling parameters	*Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*
Data quality	*Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*
Software	*Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	*Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*
Instrument	*Identify the instrument used for data collection, specifying make and model number.*
Software	*Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*
Cell population abundance	*Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*
Gating strategy	*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*
Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software — *Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*

Normalization — *If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*

Normalization template — *Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*

Noise and artifact removal — *Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*

Volume censoring — *Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

Model type and settings — *Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*

Effect(s) tested — *Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016) — *Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*

Correction — *Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity — *Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

Graph analysis — *Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

Multivariate modeling and predictive analysis — *Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*