nature food

Article

https://doi.org/10.1038/s43016-026-01361-w

Machine learning-optimized composting strategies can enhance nutrient recycling and transform food system waste into a net carbon sink

Received: 7 May 2025

Accepted: 24 April 2026

Published online: 03 June 2026

Check for updates

Lu Zhang[1,2,5], Junyu Yang[3,5], Junjie Liu[2,4,5], Qishun Zhou[3], Xuan Wang[2] ✉, Haodi Zhang[3] ✉, Yazhan Ren[2,4], Zhaohai Bai[1] & Lin Ma[1] ✉

Composting organic waste offers a circular solution to recycle nutrients and restore soil health, but nitrogen (N) and carbon (C) losses during the process undermine its agricultural and climate benefits. Here, using a machine learning approach, we analyse 848 global manure, food waste and sewage sludge composting experiments, identifying 19 key management parameters that drive ammonia (NH_3), nitrous oxide (N_2O), methane (CH_4) and carbon dioxide (CO_2) emissions. We estimate annual global losses during the composting process at 747 kt NH_3-N, 81 kt N_2O-N and 592 kt CH_4-C, equivalent to ~61 $MtCO_2e\ yr^{-1}$. By optimizing organic waste processing strategies, such as adjusting aeration, additives and C/N ratios, the composting chain could transition from a net emitter (40.1 $MtCO_2e$) to a C sink (−15.1 $MtCO_2e$), while conserving nutrients for crops. Under an optimal reduction scenario, China, Brazil and the United States emerge as the top three C sinks for composting, collectively accounting for 65% of total emission reductions.

Currently, approximately 108 Tg of nitrogen (N) is fixed annually through the Haber–Bosch process to produce fertilizer for food production[1], consuming 3–5% of global natural gas and emitting 0.46 Gt carbon dioxide equivalent per year ($CO_2e\ yr^{-1}$)[2,3]. However, only a fraction of this fixed N ultimately supports food consumption: substantial losses occur via livestock manure, sewage sludge and food waste, collectively representing 5.8 Tg of recoverable N alongside phosphorus and potassium, essential nutrients for crop production[4]. Despite their potential to substitute synthetic fertilizers, these organic wastes remain underutilized in agriculture, missing opportunities to reduce reliance on synthetic fertilizers, mitigate greenhouse gas (GHG) emissions, and advance closed-loop nutrient cycling. Closing this recycling gap would directly contribute to multiple sustainable development goals (SDGs), including SDG 2 (zero hunger) through improved soil health, SDG 3 (good health and well-being) through reduced air and water pollution, SDG 12 (responsible consumption and production) by advancing a circular economy, and SDG 13 (climate action) through the reduction of GHG emissions from waste and agriculture.

Organic waste recycling faces systemic barriers such as pathogens and antibiotics, which threaten soil and crop health if untreated wastes are directly applied to fields[5–7]. Composting emerges as a promising solution to recycle these organic wastes while mitigating

[1]State Key Laboratory of Pollution Control and Resource Reuse, School of the Environment, Nanjing University, Nanjing, China. [2]Key Laboratory of Agricultural Water Resources, Hebei Key Laboratory of Soil Ecology, Center for Agricultural Resources Research, Institute of Genetic and Developmental Biology, Chinese Academy of Sciences, Shijiazhuang, China. [3]College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China. [4]College of Advanced Agricultural Sciences, University of Chinese Academy of Sciences, Beijing, China. [5]These authors contributed equally: Lu Zhang, Junyu Yang, Junjie Liu. ✉e-mail: wangxuan@sjziam.ac.cn; hdzhang@szu.edu.cn; malin1979@nju.edu.cn

 https://doi.org/10.1038/s43016-026-01361-w

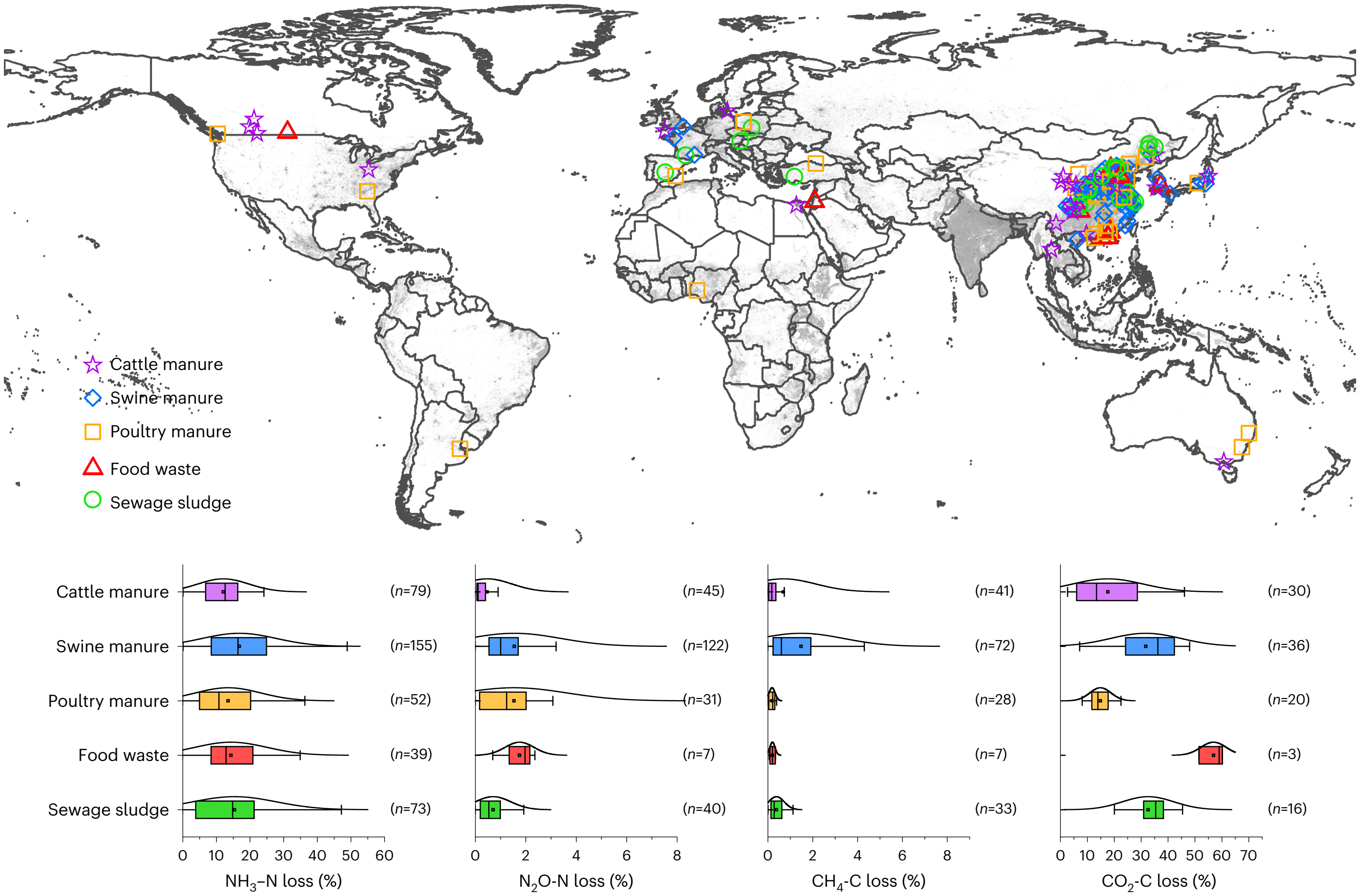

Fig. 1 | Geographical distribution of the sample sites of the global dataset and the gas loss (percentage of initial TN or TC) from different organic waste composting. The markers on the map indicate the location of the experimental sites for the composting of different raw materials. The black square represents the mean value, black solid lines in the boxplot represent the median quartiles, and box boundaries indicate upper and lower quartiles. The whiskers indicate that values extend to 1.5 orders of the box length. The values in parentheses represent the numbers of experimental observations. Basemap created with ArcGIS 10.6 with global national administrative boundary data from the Resource and Environmental Science Data Platform, IGSNRR, CAS (https://www.resdc.cn/data.aspx?DATAID=205) and Landscan population distribution data from Oak Ridge National Laboratory[53] under a Creative Commons license CC BY 4.0; Antarctica is omitted.

contamination risks[6]. When properly carried out, compost application has the potential to enhance global cereal production by 4% (96 Tg) and contribute to soil carbon (C) restoration for climate change mitigation and food security[4]. Nevertheless, approximately 30% of N and 40% of C can be lost during composting in the form of ammonia (NH_3), nitrous oxide (N_2O), methane (CH_4) and carbon dioxide (CO_2)[8,9], not only reducing nutrient recovery but also contributing to N deposition and GHG emissions that have already exceeded planetary safe operating limits[10,11].

Previous studies have attempted to quantify NH_3 and GHG emissions from composting using linear regression and meta-analysis[12,13], but such approaches struggle to capture the nonlinear, multi-gas dynamics inherent in composting. This complexity reflects the fact that N and C gas production pathways vary independently yet sometimes interact; for example, conditions that suppress NH_3 volatilization can inadvertently increase N_2O emissions[14,15]. Furthermore, the wide variety of composting materials, methods and additives further complicates prediction[8,9,13]. Emerging machine learning (ML) methods provide a powerful means to uncover such complex relationships, yet existing applications to composting are constrained by scarce datasets, restricted input variables, and limited capacity for simultaneous multi-gas prediction[16–18].

In this Article, we develop a knowledge-guided ML framework trained on 848 global observations to predict multiple pathways of N and C losses based on 19 management parameters while ensuring model predictions remain grounded in chemical and biological mechanisms[16,19]. This model was then combined with multi-objective genetic algorithms (Non-dominated Sorting Genetic Algorithm II (NSGA-II)) and life cycle assessment (LCA) to optimize composting management practices for different wastes, aiming to minimize the net environmental burdens of N losses and GHG emissions at the global level.

Results

N and C emissions under different composting managements

We collected 848 observations of N and C losses from 171 studies conducted globally between 1993 and 2023 (Fig. 1). The mean losses of NH_3-N (N mass basis) relative to initial N content ranged from 12.05% to 16.79% across organic waste types, consistent with previous estimates (9–27% of initial N)[8,20]. N_2O-N (N mass basis) losses were comparatively minor (0.47–1.75% of initial N), which is consistent with the general understanding that N_2O emissions, while climate critical, represent a minor N pathway compared to NH_3 volatilization under aerobic conditions[13,20]. CH_4-C (C mass basis) losses relative to initial total C were also low, ranging from 0.18% to 1.48%, as most reported composting trials are typically well aerated and suppress CH_4 production[9,20]. By contrast, CO_2-C (C mass basis) losses were substantially larger (14.85–56.90% of initial C), representing the primary pathway for organic C decomposition[9]. However, mass-based assessment alone is misleading

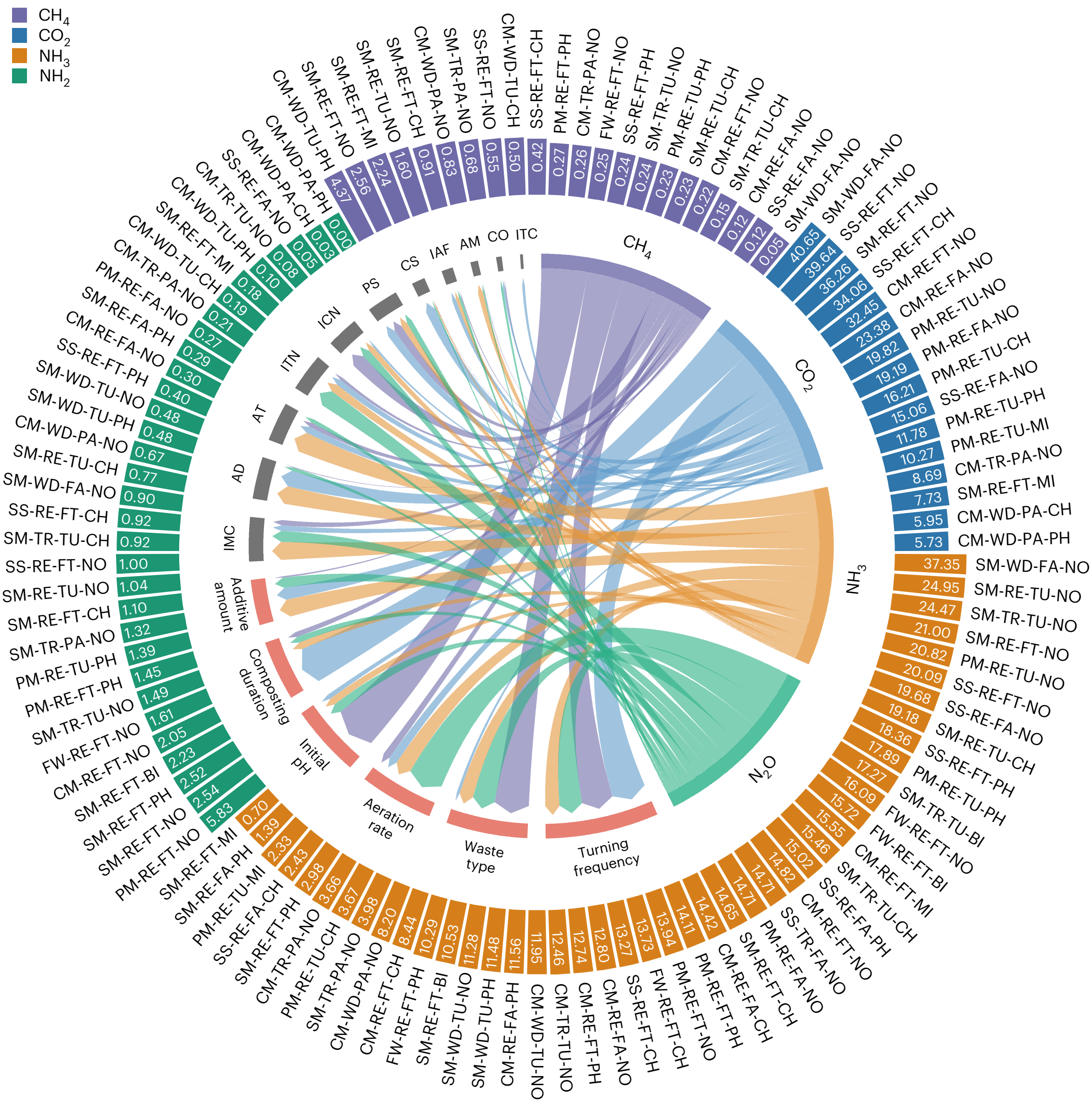

Fig. 2 | EFs conditioned on management practices and importance analysis. The outer circle shows the gas emission under different management combinations (waste type, composting system, aeration method, additive type). The numbers on the bars represent the average gas losses (percentage of initial TN or TC) during composting and are ordered from highest to lowest. Each bar must contain experimental observations ≥3. The inner circle describes the permutation importance of the features for each gas. The widths of the link lines indicate the importance of each driver for different gases. SM, swine manure; PM, poultry manure; CM, cattle manure; SS, sewage sludge; FW, food waste; RE, reactor; TR, trough; WD, windrow; FT, forced aeration + turning; FA, forced aeration only; TU, turning only; PA, passive aeration; PH, physical additive; CH, chemical additive; BI, biological additive; MI, mixture of the above additives (two or more); NO, no additive; IMC, initial moisture content; ITC, initial total carbon; ICN, initial C/N; CS, composting systems; AM, aeration method; IAF, intermittent aeration frequency; AD, aeration duration; PS, pile size; AT, additive type.

for climate impact. When converted to CO_2e using 100 year global warming potentials (GWP-100), N_2O and CH_4 together contributed over 90% of the total GHG footprint (Supplementary Fig. 5), underscoring the critical need to target non-CO_2 gases mitigation.

N and C losses during composting varied considerably among waste types (Fig. 1), driven primarily by differences in initial biochemical composition[8,13]. Cattle manure, which typically has higher C/N ratio and lignocellulose content[13], showed lower CO_2 (17.61%) and NH_3 (12.05%) losses, whereas swine and poultry manures lost more NH_3-N (13.47–16.79%) owing to higher mineralizable N content[20]. Food waste composting showed higher N_2O-N (1.75%) losses, primarily driven by its high moisture content and poor porosity[13]. Meanwhile, sewage sludge showed moderate C (32.59% CO_2-C, 0.38% CH_4-C) and N (15.38% NH_3-N, 0.70% N_2O-N) losses, influenced by its diverse sources and various pre-treatment processes[6,21]. The N and C losses are more closely related to the differences in management practices (Fig. 2). For example, the highest NH_3-N loss was observed with the combination of swine manure–windrow–forced aeration only–no additive (SM–WD–FA–NO) treatment, accounting for 37.35% of total initial N, primarily owing to the open, forced aeration composting system and the higher initial content of easily mineralizable N (ref. 13). By contrast, adding different additives during the composting can greatly reduce the NH_3-N losses to 0.70–1.39%, such as in the swine manure–reactor–forced aeration + turning–mixture of additives (SM–RE–FT–MI) and swine manure–reactor–forced aeration only–physical additive (SM–RE–FA–PH) treatments (Fig. 2). This reduction is primarily a result of adsorption and/or fixation of ammonium N by additives[20].

The complexity of the production mechanism results in a wide range of N_2O-N loss during composting, with losses highest at 5.83% observed in the poultry manure–reactor–forced aeration + turning–no

https://doi.org/10.1038/s43016-026-01361-w

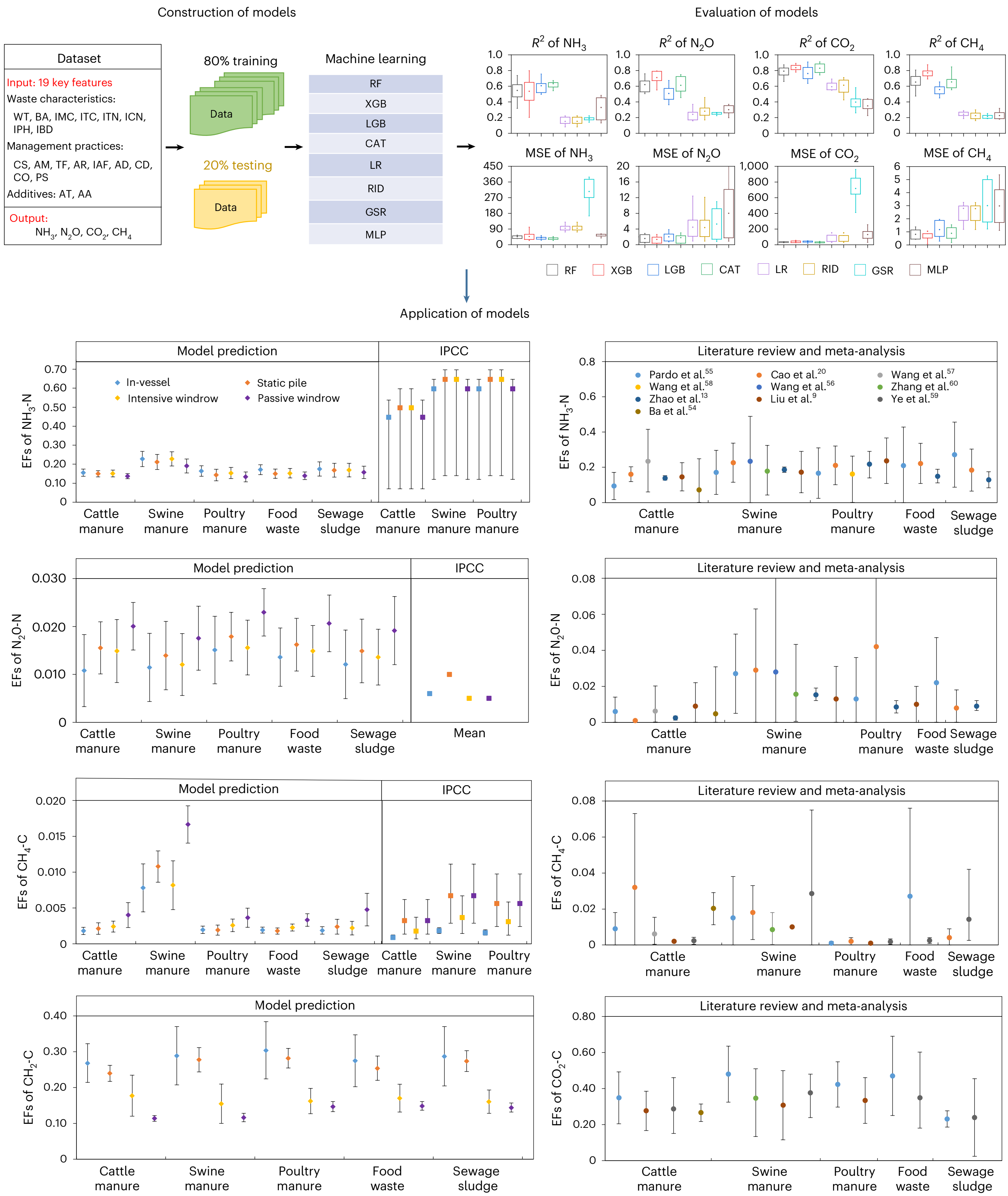

Fig. 3 | The framework for the construction, evaluation and application of the composting simulation model. A larger R^2 and smaller m.s.e. indicate higher accuracy in model predictions. R^2 is dimensionless; m.s.e., calculated from percentage-scale response variables, is expressed in $\%^2$; EF is shown as a fraction of initial TN or TC. The centre point indicates the mean, and the box boundaries indicate the upper and lower quartiles. WT, waste type; BA, bulking agent type; IPH, initial pH; IBD, initial bulk density; TF, turning frequency; AR, aeration rate; CD, composting duration; CO, covered; AA, additive amount. RF, random forest; XGB, eXtreme Gradient Boosting; LGB, Light Gradient Boosting; CAT, Categorical Boosting; LR, linear regression; RID, ridge regression; GSR, Gaussian process regression; MLP, multilayer perceptron.

https://doi.org/10.1038/s43016-026-01361-w

Material characteristics					Management							Additives		Gas emissions (%)				Net C footprint (CO_2eq t^{-1})
Waste Type	Bulking agent	Initial C/N	Initial moisture (%)	Initial pH	System	Aeration method	Aeration rate (l min^{-1})	IAF (h/h)	Aeration duration (days)	Turning frequency (day/time)	Duration (days)	Type	Amount (%DM)	NH_3-N	N_2O-N	CH_4-C	CO_2-C	
Cattle manure	Green waste	30	56	7.7	Reactor	Intermittent	0.03	1	30	-	58	Gypsum	4	5.00	0.04	0.09	16.58	−61
		33	58	6.6	Reactor	Continuous	0.10	-	40	-	40	Inorganic acid	10	1.03	0.10	0.18	18.54	128
Swine manure	Green waste	20	53	6.4	Windrow	Intermittent	0.85	3	36	2	67	Clay	4	7.03	0.28	0.01	14.07	−31
		28	60	6.3	Reactor	Intermittent	0.16	2	41	5	41	Inorganic acid	9	0.70	0.99	2.71	24.05	550
Poultry manure	Sawdust	34	57	6.7	Reactor	Intermittent	0.71	3	25	-	29	Gypsum	5	10.77	0.08	0.09	13.85	−71
		30	70	5.5	Reactor	Continuous	0.32	-	45	-	49	Ca-superphosphate	6	1.10	0.63	0.05	12.25	14
Food waste	Green waste	31	56	7.9	Reactor	Continuous	0.12	-	19	-	57	Clay	6	7.60	0.15	0.05	16.93	−53
		30	50	6.6	Reactor	Continuous	0.11	-	37	-	53	Gypsum	9	1.55	0.48	0.05	19.01	−31
Sewage sludge	Green waste	32	69	7.8	Reactor	Intermittent	0.11	1	26	-	44	Mature compost	4	6.00	0.12	0.09	14.82	−56
		33	55	6.4	Reactor	Intermittent	0.13	1	40	-	55	Ca-Mg-P fertilizer	10	1.09	0.96	0.13	11.73	144

Fig. 4 | Tailored composting management strategies for optimal GHG and reactive N mitigation, respectively. The blue background indicates the optimal strategy for reducing reactive N emissions, while the white background represents the optimal strategy for reducing GHGs. h/h, hours of aeration per hour of operation; day/time, interval in days between turning events; %DM, percentage of dry matter. Figure created in BioRender; Liu, J. https://BioRender.com/k8q5j1l. (2026).

additive (PM−RE−FT−NO) treatment (Fig. 2). Additives and aeration offer greater potential for N_2O-N control. The turning only and forced aeration only practices typically result in lower N_2O-N losses compared to the forced aeration + turning treatment. This is because the combined aeration and turning create a dynamically fluctuating redox environment, creating ideal conditions for the co-occurrence and coupling of nitrification (aerobic) and denitrification (anaerobic), which is a key driver of N_2O emissions[14,22]. Similar to the control of NH_3-N losses, additives can effectively reduce N_2O-N losses by approximately 12.2% to 75.2% (Fig. 2). However, it is important to note that the types of additives effective in reducing NH_3 volatilization (for example, acidifying agents) often differ from those that suppress N_2O (for example, nitrification inhibitors), as NH_3 emission is primarily governed by physico-chemical conditions whereas N_2O production is driven by microbial N transformation processes[9,20,22]. Such mechanistic divergence explains why conventional single-approach strategies often fail to reduce both gases simultaneously[14,20,21], highlighting the need for integrated, multi-target optimization.

As a microbe-dominated aerobic degradation process, CO_2 emissions are unavoidable during composting[23]. CO_2-C loss can account for 5.73% of initial C content in the composted material and up to 40.65% under different composting combinations (Fig. 2). Passive aeration generally results in lower CO_2 emissions but may prolong the composting duration and result in incomplete maturation[6]. A certain amount of CO_2 emissions is crucial, as it indicates the microbial degradation of toxic substances, which is essential for producing crop-beneficial compost[6]. Simultaneously, minimizing unnecessary CO_2 loss while ensuring maturation improves the C sequestration potential of the final compost product[23]. CH_4 is mainly produced by methanogens in local anaerobic environments within the compost pile, ranging from 0.05% to 4.37% of initial C content under different management practices (Fig. 2). Sufficient diffusion of oxygen into the matrix can suppress the activity of anaerobic methanogens while concurrently stimulating methanotrophic oxidation of CH_4 to CO_2 (ref. 24).

Key factors determining gas emissions during the composting process

The key factors affecting the different N and C losses during composting are varied (Fig. 2). Turning frequency emerged as the dominant factor for both N and C losses. Regular turning optimizes oxygen diffusion and enhances material homogeneity, facilitating organic matter mineralization while inhibiting CH_4 generation[8,9]. However, turning may simultaneously promote NH_3 volatilization and drive the redistribution of NO_x^- from aerobic surface zones to internal anaerobic regions, facilitating incomplete nitrification–denitrification and increasing N_2O emissions[9,22]. Waste type also plays a crucial role (Fig. 2), primarily owing to the specific physicochemical properties and indigenous microbial communities that govern C and N transformation pathways[9,13].

Apart from these factors, additive application and aeration parameters emerged as critical drivers of NH_3 loss, including additive type, additive amount, aeration duration, and aeration rate (Fig. 2). Specifically, additives effectively stabilize ammonium N through chemical immobilization or physical adsorption, thereby mitigating NH_3 emissions[20], whereas longer aeration duration and higher aeration rate typically enhance NH_3 volatilization by accelerating organic N mineralization and promoting ammonium N diffusion from the compost matrix[9]. For N_2O, aeration rate and initial total N (ITN) were identified as the most important factors ($P < 0.01$) after turning frequency and waste type (Supplementary Fig. 8). Higher aeration inhibits N_2O reductase activity and enhancing nitrification, thereby promoting N_2O emissions[13,22]. Concurrently, substrate N enrichment (elevated NH_4^+-N and NO_3^--N concentrations) promotes dual nitrification–denitrification pathways, further increasing N_2O emissions[14,25].

Composting duration and aeration duration significantly affect CO_2 emissions ($P < 0.01$) (Fig. 2 and Supplementary Figs. 7c and 8c), as prolonged aerobic conditions inevitably lead to excessive mineralization of organic C (ref. 9). pH critically regulates CH_4 emissions, with acidic (<6.0) or alkaline (>8.5) conditions suppressing methanogen metabolic activity[24]. Overall, these results highlight the complexity of gas emission factors (EFs) across composting management practices, underscoring the need for comprehensive models that integrate multiple management variables to accurately predict and mitigate emissions.

Machine learning enabled prediction of N and C losses from composting

We used 8 ML algorithms to predict the EFs for each of the four gases based on 19 management parameters (Fig. 3). Categorical Boosting

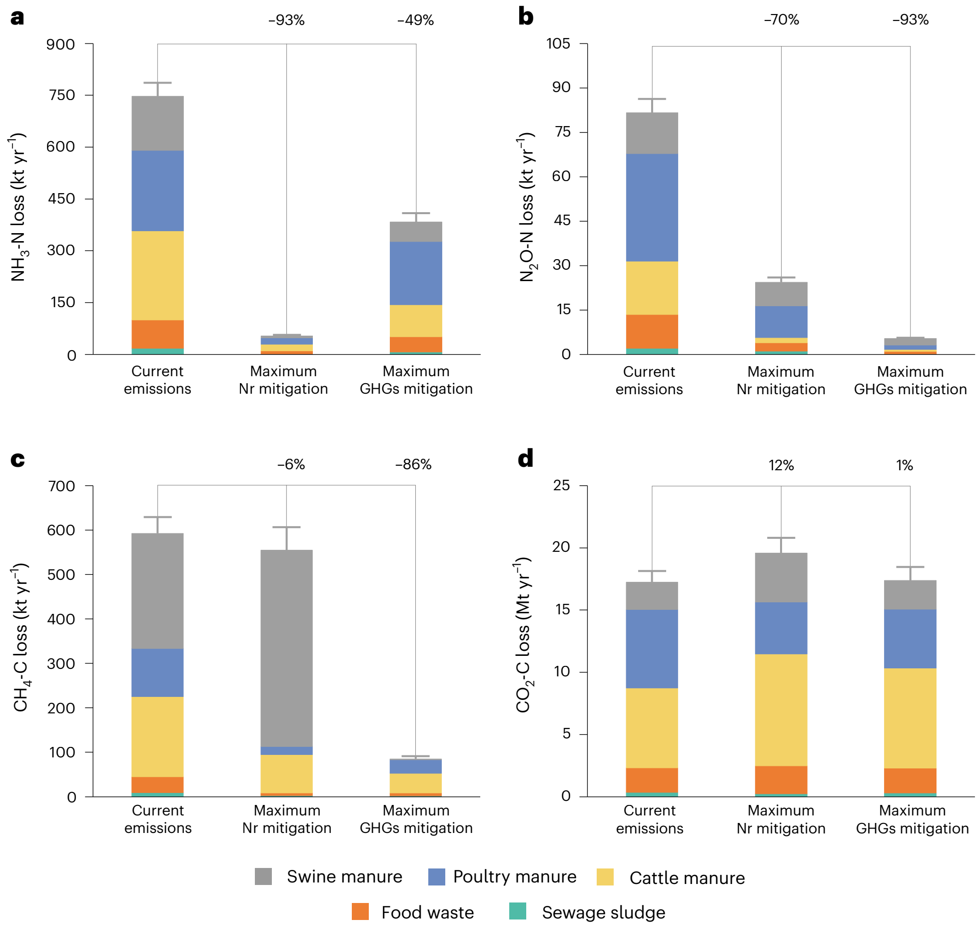

Fig. 5 | Changes in C and N gas emissions from global organic waste composting with the adoption of optimized management strategies. **a**–**d**, NH_3-N (**a**), N_2O-N (**b**), CH_4-C (**c**) and CO_2-C (**d**) emissions from composting under current practices, Nr mitigation strategies and GHGs mitigation strategies. Data are presented as mean ± s.d. n = 10,000 independent Monte Carlo iterations per strategy. Error bars represent ±1 s.d. calculated across the 10,000 independent simulations per group, reflecting parametric uncertainty. All statistical summaries strictly reflect independent computational variability.

achieved the best performance for NH_3 (R^2 = 0.63 ± 0.07), while eXtreme Gradient Boosting was optimal for N_2O (R^2 = 0.71 ± 0.07), CO_2 (R^2 = 0.84 ± 0.03) and CH_4 (R^2 = 0.77 ± 0.05). While previous ML studies have reported higher predictive accuracy (often R^2 > 0.9) for composting GHG emissions[16–18], these models were typically trained on specific waste types or controlled composting scenarios with more homogeneous datasets (Supplementary Table 8).

By contrast, our model was developed on a globally representative dataset encompassing 19 operational parameters, such as different types of waste and composting methods, prioritizing generalizability over narrow predictive precision. The moderate R^2 values thus reflect the inherent variability of real-world composting rather than model inadequacy.

The ML-derived EFs revealed notable discrepancies with current IPCC (Intergovernmental Panel on Climate Change) defaults (Fig. 3). The IPCC (2019) guidelines provide default EFs for four different composting management systems (in-vessel, static pile, intensive windrow, passive windrow), reporting combined NH_3 and NO_x losses as 45–65% of initial N at Tier 1 (ref. 26). Considering NO_x-N emissions from composting are typically less than 1% of the initial N (ref. 27), this default range effectively represents NH_3 emissions. Our ML model shows that EF of NH_3 ranges from 13% to 23% of initial N, with minor variation observed across different composting methods and materials. This prediction aligns with current meta-analysis and literature review (Fig. 3), indicating that IPCC default value may overestimate NH_3 emission from composting. Conversely, N_2O EFs predicted by ML (0.5–2.3% of initial N) were approximately twice the IPCC Tier 1 default for most waste types except cattle manure. This discrepancy was also observed in other studies, suggesting that the IPCC underestimated N_2O emissions from composting[8,13]. N_2O EFs varied substantially across waste types, suggesting that material-specific EFs are needed for accurate emission assessments. CH_4 EFs were broadly consistent with IPCC Tier 2 (taking into account manure characteristics and management systems) estimates (0.1–1.7% of initial C). For CO_2, where IPCC provides no default EF, as this process is considered part of natural C cycling, our model estimated losses of 10–30% of initial C, offering a basis for evaluating the C sequestration potential of composting.

N and C losses from composting at the global level

Composting is widely used as an effective technology for recovering organic resources from waste, particularly for manure treatment in China, the United States and Brazil (Supplementary Figs. 17–19). In Europe, manure is rarely composted, whereas a high proportion of food waste and sludge is composted (Supplementary Fig. 18). By using the updated EFs, we estimated that 747 ± 38 kt of NH_3-N, 81 ± 4 kt of N_2O-N and 592 ± 36 kt of CH_4-C were emitted annually from global composting in 2020 (Fig. 5). These account for 6.4% of NH_3 emissions, 23.8% of direct N_2O emissions and 5.5% of CH_4 emissions from global manure management[28,29], highlighting that composting is an important contributor and will increase with agricultural intensification and expanding municipal organic waste collection. Cattle and poultry manure composting dominated NH_3, N_2O and CO_2 emissions owing to their large processed volumes (Supplementary Fig. 18), whereas swine manure contributed disproportionately to CH_4 emissions (24%) owing to its higher CH_4 EF.

 https://doi.org/10.1038/s43016-026-01361-w

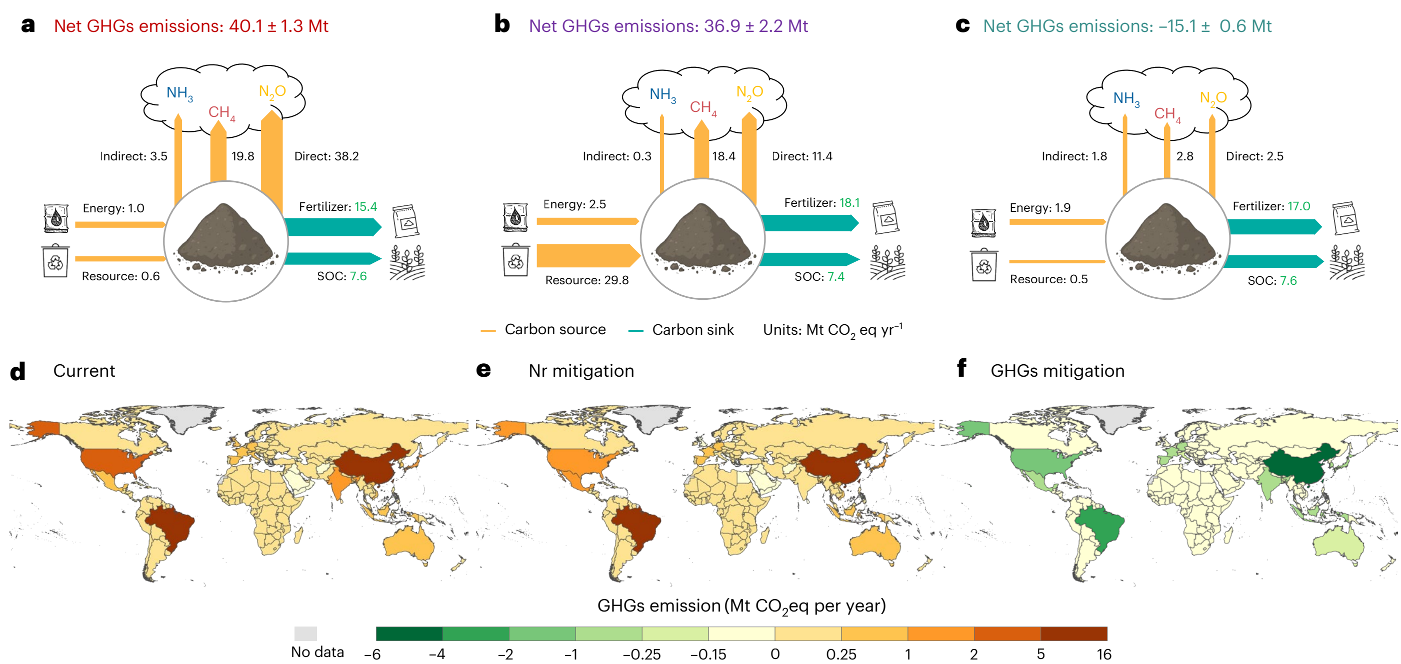

Fig. 6 | Global GHGs emissions from whole organic waste composting chain by adopting integrated management strategies. a–f, Current scenario (**a**,**d**), Nr mitigation scenarios (**b**,**e**) and GHGs mitigation scenarios (**c**,**f**). Energy, the fuel and electricity used for waste collection, transport and composting operations (aeration, turning); Resource, the C emissions from all material inputs to the composting process, including composting facility and additives; Fertilizer substitution benefit, avoided emissions from synthetic N fertilizer production as a result of compost use; SOC benefit, C sequestered in soil through compost application. Basemaps in **d**–**f** created with ArcGIS 10.6 with global national administrative boundary data from the Resource and Environmental Science Data Platform, IGSNRR, CAS (https://www.resdc.cn/data.aspx?DATAID=205); Antarctica is omitted. Schematics created in BioRender; Liu, J. https://BioRender.com/237tjfj (2026).

The predicted values are different from those of IPCC, especially for NH_3 and N_2O emissions, when considering the more accurate EFs of specific composting materials (Supplementary Fig. 21). Specifically, global NH_3 volatilization from manure composting was 73% lower (equal to 1,723 kt NH_3-N) than IPCC-derived estimates, while N_2O emission was 171% higher (equal to 42 kt N_2O-N yr^{-1}), reflecting the importance of waste-specific parameterization. Large country-level variations in N_2O emissions were also observed, particularly in the United States, Canada and Brazil, driven by their high share of swine and poultry manure, wastes with N_2O EFs far above the IPCC default (Supplementary Table 4).

Machine-learning-guided strategies for lower reactive N losses and net C footprint

Using our ML model coupled with NSGA-II multi-objective optimization, we generated 100,000 optimal composting management combinations (20,000 per waste type) to identify strategies that minimize N and C losses while maintaining compost quality (germination index > 80%) (Fig. 4 and Supplementary Fig. 11). The resulting Pareto fronts reveal inherent trade-offs between emission pathways; targeting a single gas does not yield the lowest net C footprint owing to inter-gas trade-offs and indirect emissions from management practices themselves (Supplementary Figs. 12–14 and 22). Here we present two representative scenarios: (1) minimizing reactive N (Nr) losses (NH_3-N + N_2O-N) and (2) minimizing net C footprint, which accounts for the direct GHG emissions (N_2O, CH_4) and indirect emissions from composting operations (energy consumption, additive consumption and so on), avoiding emissions through substituting synthetic N fertilizer and the long-term soil C sequestration via properly recycling compost to crop production (Fig. 6 and Supplementary Fig. 20).

In general, minimizing NH_3 volatilization is prioritized to reduce N losses during composting. Our results suggest that the optimized Nr reduction strategy integrates three key interventions: (1) maintaining initial C/N ratios at 28–33 by supplementing with lignocellulosic biomass (for example, green waste, sawdust) to enhance NH_4^+ immobilization via microbial assimilation; (2) incorporating 6–10% acidic additives (for example, gypsum, inorganic acids, Ca-superphosphate) as pH buffer to shift the NH_3/NH_4^+ equilibrium; and (3) implementing low-frequency turning combined with low aeration rate (0.1–0.3 l kg^{-1} min^{-1}) to minimize NH_3 release from compost pile (Fig. 4). By contrast, the low-C emission strategy prioritizes low-footprint additives (for example, clay, mature compost) and intermittent aeration to suppress N_2O and CH_4 without increasing energy consumption (Fig. 4).

Under the Nr minimization scenario, our results suggest that global composting could reduce N losses by 694 kt NH_3-N and 57 kt N_2O-N (70–93% below current levels; Fig. 5a,b). The conserved N, equivalent to 0.7% of global N fertilizer use, would reduce economic and environmental costs of fertilizer production[3]. This scenario also delivers a 6% reduction in CH_4-C but a 12% increase in CO_2-C emissions (Fig. 5c,d), reflecting the trade-off between C and N reductions. By contrast, the lowest C footprint scenario achieved only a 49% reduction in NH_3-N losses but 93% and 86% reductions in N_2O-N and CH_4-C losses, respectively, with a smaller CO_2-C increase (Fig. 5). These synergistic reductions stem from the integration of low-C-footprint additives and optimized aeration (Fig. 4).

From a life cycle perspective, system boundaries encompass the following: (1) upstream processes, including organic waste collection and transport; (2) the composting process, encompassing direct GHG and NH_3 emissions and indirect emissions from energy and additive consumption; and (3) downstream impacts, accounting for avoided emissions from synthetic fertilizer substitution and long-term soil C sequestration following compost application (Supplementary Fig. 20). The net GHG emissions were estimated at 40.1 $MtCO_2e$ from global organic waste composting in 2020 (Fig. 6a), which were primarily attributed to the direct N_2O emissions (38.2 $MtCO_2e$) and CH_4 emissions (19.8 $MtCO_2e$), then followed by indirect N_2O emission via NH_3 emission (3.5 $MtCO_2e$) and the energy use in composting (1.0 $MtCO_2e$) in 2020. These emissions could be offset by around 15.4 $MtCO_2e$ through the replacement of synthetic N fertilizer and by 7.6 $MtCO_2e$ through the soil C storage, considering the overall C input rate and the 20 years long-term decomposition process[30].

Under the Nr minimization scenario, net emissions would decrease slightly to 36.9 $MtCO_2e$ (Fig. 6b). Although direct and indirect N_2O emissions drop substantially (to 11.4 and 0.3 $MtCO_2e$, respectively), the required inputs of gypsum and acids raise resource-related emissions from 0.6 to 29.8 $MtCO_2e$. This solution yields a further reduction of 2.7 $MtCO_2e$ through more N conserved in the final commercial composting fertilizer. By contrast, the lowest C footprint scenario transforms the system from a C source to a sink, with net emissions of −15.1 $MtCO_2e$ (Fig. 6c), driven by substantial reductions in direct and indirect N_2O (−37.4 $MtCO_2e$) and CH_4 (−17.0 $MtCO_2e$) without a substantial increase in resource-related emissions.

Overall, optimal composting management can shift the global composting chain from a net emitter to a C sink, reducing annual emissions by 55.2 $MtCO_2e$ (Fig. 6). China, Brazil and the United States emerge as the top three C sinks, collectively accounting for 65% of total emission reductions (Fig. 6d–f and Supplementary Fig. 25). These mitigation strategies also deliver co-benefits for terrestrial ecosystems, including avoided acidification equivalent to 1,077.2 kt SO_2 per year (Supplementary Fig. 26), underscoring the value of best-practice composting in closing nutrient loops and supporting biodiversity.

Discussion

Our study indicates that waste composting has substantial C sink potential when managed effectively. Although the proportion of waste currently being composted is relatively low, the recycling of nutrient resources is actively promoted worldwide, with composting rates increasing annually[31]. For example, Ireland has set a target of composting 70% of organic waste by 2030 (ref. 32). Our ML-guided framework can serve as a scientific tool for advancing waste composting management, assisting farmers and compost producers in designing optimized, context-specific strategies. However, it is important to note that the optimization strategies were designed based on industrial scale to better guide production practices. While trained on multi-scale data (from lab scale to industrial scale), the model can capture fundamental relationships that are largely scale invariant, providing robust guidance on key management principles (for example, optimal C/N ratio). Nevertheless, the actual efficacy of these strategies in controlling gaseous emissions may vary across composting scales because of differences in physical conditions and operational implementation (Supplementary Fig. 24). Future work should integrate engineering principles to build scale-explicit prediction frameworks and, with high-resolution input data, enable time-series modelling for dynamic process control.

Beyond the two scenarios presented here, our framework can generate differentiated strategies tailored to region-specific priorities. For example, to meet the 30% CH_4 reduction target between 2020 and 2030, equivalent to 18 Mt CO_2e yr^{-1} in solid waste management[33,34]. Our study showed that composting systems could deliver comparable mitigation if designed to minimize CH_4 emission (Supplementary Figs. 14 and 20). Similarly, in regions facing severe NH_3-related biodiversity and air pollution risks, the lowest NH_3 emission mitigation solution in composting could also reduce NH_3-N emission by 92% (Supplementary Figs. 12 and 20), which could also help to alleviate the negative effects of NH_3 at the regional level. Furthermore, beyond pursuing minimizing either GHG emissions or Nr loss alone, compromise solutions can also be designed to address multiple objectives simultaneously, such as maximizing N conservation while ensuring the system acts as a net C sink (net GHG emissions < 0) or minimizing the net C footprint while ensuring Nr reduction exceeds 80% (Supplementary Figs. 15, 16 and 23). Such trade-off solutions highlight the model's utility in supporting context-specific decision-making and can be used to guide the achievement of sustainable ecosystems at the country and local levels.

This study used a data-driven approach to address the nonlinear effects and complex interactions between C and N losses with waste properties and management strategies, which was previously impossible with process- and experience-based approaches. However, further improvements are needed to address model uncertainties in the future: (1) expanding the dataset with more in situ observations and incorporating additional drivers such as climate and season and (2) integrating data-driven ML and biophysical-based models to develop hybrid models, while their configuration and parameters can be optimized through an ensemble learning approach[19,35]. Such advances will strengthen the reliability and transferability of composting optimization across diverse environmental and operational contexts, ultimately supporting the transition toward closed-loop nutrient cycling and climate-resilient food systems.

Methods

Data collection

For this study, a comprehensive literature review was conducted using academic databases, such as Web of Science, Google Scholar and the China National Knowledge Infrastructure. The search covered the period from 1993 to 2023 to ensure that comprehensive composting research was included at the outset of the project. The keywords used in the search included the following: composting, ammonia or NH_3, nitrous oxide or N_2O, greenhouse gases or GHGs, methane or CH_4, carbon dioxide or CO_2, nitrogen loss or N loss, and C loss or carbon loss.

To be included in this database, studies had to meet the following criteria: (1) the composting process must be complete; (2) laboratory-scale incubation experiments are not included; and (3) the study must include emission factors for at least one of NH_3, N_2O, CO_2 and CH_4. As a result, 171 papers with 848 observations were eventually included in the analyses. To extract specific data from figures in the selected literature, GetData Graph Digitizer 2.22 was used. All relevant primary trials and meta-analyses were included in the dataset (Supplementary Reference Lists 1 and 2). To accurately and comprehensively establish the relationship between gas emission characteristics and management parameters, the dataset consisted of 19 key features, including waste characteristics (waste type, bulking agent type, initial moisture content, initial total carbon, ITN, initial C/N, initial pH and initial bulk density), management practices (composting systems, aeration method, turning frequency, aeration rate, intermittent aeration frequency, aeration duration, composting duration, covered and pile size) and additives use (additive type and amount) (Supplementary Table 1 and Supplementary Fig. 1). In addition, the germination index of compost products was also included in the database and used as a predictor, considering that compost may be intended for use as organic fertilizers[36]. Data processing and visualization were performed using a combination of Python 3.8, Microsoft Excel 2024, OriginLab 2026 and ArcGIS 10.6.

Data preprocessing

A series of data preprocessing steps was performed to make the dataset more suitable for modelling. First, categorical features (waste type, bulking agent type, composting systems, additive type and so on) were digitized using sequence coding. Second, a local outlier factor analysis was used to assess the local relative density of each sample point and identify outliers for exclusion[37,38]. Then, we calculated the Pearson correlation matrix to identify and address potential multicollinearity. The absolute value of Pearson's correlation coefficient was between 0.1 and 0.4 (Supplementary Fig. 2), indicating that no strong collinearity overlap existed between variables[16]. Finally, we processed the missing values of the input features for different prediction objectives, where input features with more than 25% missing values were eliminated, whereas input features with a small number of missing values were filled using K-nearest neighbours estimation[39,40].

Knowledge-guided ML model development and evaluation

First, the selection of 19 input variables was guided by prior knowledge of composting biochemistry to ensure mechanistic relevance. To ensure the biological and chemical plausibility of model predictions, we imposed physicochemical bounds on key input variables. These constraints, derived from established composting literature, prevented the model from extrapolating into unrealistic parameter spaces. For instance, the pH of composting materials was bounded within a realistic range (for example, 5.0 to 9.5), as values below 5.0 would strongly inhibit most microbial activity and are not sustained in typical composting systems[6]. The original dataset was randomly partitioned 8:2 into training and test sets. Within the training set, we performed K-fold cross-validation ($K = 10$) for model training. Eight different models were used for target prediction, including random forest, eXtreme Gradient Boosting, Light Gradient Boosting Machine, Categorical Boosting, linear regression, ridge regression, Gaussian process regression, and multilayer perceptron. To enhance predictive performance, Bayesian optimization was used for hyperparameter tuning, effectively exploring the hyperparameter space to identify optimal model parameter combinations[41]. The root mean square error and coefficient of determination (R^2) values generated on the test set were used as metrics to compare the predictive performance of the models[41,42]. To reduce random errors and enhance generalization, we trained 10 random splits of the dataset and used the average performance as the final criterion for model selection. Finally, the best model selected for each prediction target was finally trained on the entire dataset to ensure the development of a robust final model. All model training processes were performed using Python 3.8.

Feature importance and interaction analysis

Feature importance analysis plays a critical role in improving the model interpretability[43,44]. In our study, we used multiple metrics to assess the feature importance based on our optimal ML model, including mean square error (m.s.e.) increase, node purity increase, P value and Shapley additive explanations value, which helps mitigate the limitations of relying on a single metric[45]. Features with higher Shapley additive explanations values, larger m.s.e., and greater node purity values generally contributed more strongly to the model outputs[44,45].

Prediction of gas EFs from various wastes under different composting management systems

To establish baseline EFs, we used our optimal ML model to predict C and N gas EFs for different waste types under the four composting systems defined by the IPCC (2019 Refinement, Chapter 10, Table 10.21) (Supplementary Table 2). These predictions represent a baseline scenario, assuming no advanced control measures (for example, additives, aeration optimization). We then performed a systematic comparison of these ML-predicted EFs with the corresponding default EFs from the IPCC guidelines (Tier 1 for NH_3 and N_2O, Tier 2 for CH_4). To quantify prediction uncertainty, for each EF we conducted 500 Monte Carlo simulations by randomly sampling key input features (for example, ITN, C/N ratio) within their observed ranges in the database and reported the mean and standard deviation.

Co-optimization of multiple gas emission reductions

Based on the optimal machine learning model, we performed multi-objective optimization using the NSGA-II, an efficient genetic algorithm that excels at classifying and maintaining elite solutions and is designed to search for the Pareto-optimal frontier, which represents the optimal trade-off between competing objectives[46]. As the compost product must be plant friendly to serve as organic fertilizer, the germination index of the compost must exceed 80% (ref. 36). To further guide practical production, the composting scale (pile size) was set at industrial scale (>10 m^3). Management practices (bulking agent type, initial C/N, initial pH, initial moisture content, composting systems, aeration method, turning frequency, aeration rate, intermittent aeration frequency, composting duration, covered, additive type, additive amount) were then used as adjustable parameters to optimize the objective function, with the aim of minimizing emissions of multiple gases during composting and achieving a final germination index > 80%. Initially, the parameter boundaries for management practices were defined based on the database values. The Pareto solution set was then obtained by comparing the objective function values through 100,000 iterative runs of NSGA-II.

Considering that changes in management practices have an indirect environmental footprint, a LCA methodology was used to estimate the net environmental footprint of the different management practices based on obtaining the Pareto solution set. The system boundary encompasses three key stages: (1) upstream (collection and transport of waste), (2) the composting process itself (direct and operational emissions) and (3) downstream (agronomic benefits from compost use, including fertilizer substitution and soil C sequestration) (Supplementary Fig. 20). The LCA in this study is based on the ReCiPe impact assessment technique extracted from the Eco-invent database, which converts a list of full life-cycle inventory results into a limited number of indicator scores, including midpoint indicators and a normalized total environmental impact indicator[47,48]. We evaluated the potential environmental impacts of various composting management practices using the treatment of 1 t of waste (dry weight basis) as the functional unit. The environmental footprints of various composting management practices are shown in the Supplementary Table 6. CO_2 emissions from the composting process were considered to have no net climate impact. However, the reduced CO_2 could eventually form stable soil organic C (accounting for about 8%), thereby acting as a C sink[4,30]. In addition, the reduced N loss can replace N fertilizer thus reducing the environmental footprint of N fertilizer production. The preparation and transport of materials were also considered in this study as we assumed that these factors were the same for all composting processes. The formulas for the assessment of the C footprint of the different composting management practices are as follows:

$$C_{Total} = C_{tech} + C_{gas} - C_{Fertilizer} - C_{SOC} \tag{1}$$

where C_{Total} is the total C footprint (kg CO_2e per tonne dry matter); C_{tech} is the total C footprint of the different composting technologies (kg CO_2e per tonne dry matter), including the preparation and transport of materials, composting systems, aeration systems, additives and so on (Supplementary Fig. 20 and Supplementary Table 6); C_{gas} is the C footprint of gaseous emissions from the composting process (kg CO_2e per tonne dry matter), including NH_3, N_2O and CH_4. The conversion factor for NH_3-N to N_2O-N is 0.01, while the coefficients for converting N_2O and CH_4 into CO_2 equivalents are 298 and 25, respectively; $C_{Fertilizer}$ is the C footprint of replacing N fertilizers (kg CO_2e per tonne dry matter); and C_{soc} is the environmental footprint of C sequestration (accounting for about 8%) in soils (kg CO_2e per tonne dry matter).

The mitigation potential for each GHG was calculated as the difference between emissions under current (baseline) practices and those under optimized management scenarios.

Global mitigation potential assessment

We conducted a comprehensive assessment of the global potential for reducing emissions of reactive N and GHG emissions through model-optimized composting management practices (Supplementary Fig. 10). Specifically, three types of manure (cattle, poultry and swine), sewage sludge and food waste were included in the global analysis as the main composting feedstocks. In the first step, we estimated the annual global production of the five target organic wastes. The generation of animal manure was sourced from the FAO database[29], while data for food waste and sewage sludge were drawn from established literature[33,49–51]. The total potential C and N quantities from each waste stream were then calculated by their respective C–N content and ratios (Supplementary Fig. 4). In the second step, we quantified the fraction of produced waste that is actually composted. Country- or region-specific composting fraction (that is, the proportion of collected waste used for composting) for each waste type were compiled from the literature[13,33,50–52]. The amounts of C and N used for composting were subsequently calculated by multiplying the total quantities by these composting fractions. All detailed data sources are provided in Supplementary Table 7 and Supplementary Data 1. In the third step, we calculated the total global gas emissions from the composting process based on the proportion of all current waste composting management systems and the gas emission factors of the different wastes (Supplementary Table 4). Finally, the global emission reductions of different gases were simultaneously calculated based on the gas EFs of the optimized composting management practices. All map-related operations were performed using ArcGIS 10.6 software.

Uncertainties

For global predictions, uncertainty was associated with both the accuracy of machine learning model predictions and the quality of data collected from global and regional statistics, as most of these studies were carried out in China, Europe and Japan (Fig. 1). For example, the datasets for N_2O, CH_4 and CO_2 had fewer missing values, leading to more favourable prediction outcomes, when compared to NH_3 emissions. In addition, some other key factors influencing gas emissions during composting (for example, climate, season) were not included because of insufficient data reports, which could affect the predictive performance of the model[8]. To address uncertainty in model predictions, we conducted 200 random runs for each EF using bootstrap and stratified sampling methods, then analysed the model outputs by calculating the means, standard deviations and confidence intervals[45]. The relative uncertainty of the data from global and regional statistics, surveys and other sources (expressed as the coefficient of variation) was estimated based on published literature and the authors' expert judgment[33,51]. Furthermore, we performed a Monte Carlo simulation with 10,000 runs to characterize the total uncertainty in global C and N gas emissions from composting, based on variations in global input data and EFs following uniform and normal distributions, respectively.

Statistics and reproducibility

No statistical method was used to predetermine sample size; the dataset size (848 observations from 171 studies) was determined by literature availability meeting our inclusion criteria. Laboratory-scale incubation experiments were excluded at data collection, and statistical outliers were removed via local outlier factor analysis (Supplementary Table 2); no other data were excluded from the analyses. The experiments were not randomized, as this is a retrospective meta-analysis; model training used random 80:20 data partitioning with 10-fold cross-validation. The investigators were not blinded to allocation during outcome assessment; automated pipelines with pre-defined metrics (R^2, m.s.e.) minimized subjective bias.

Reporting summary

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

Data availability

The datasets supporting the findings of this study are publicly accessible. The core dataset required to execute the main program has been permanently archived in Zenodo and is available at https://doi.org/10.5281/zenodo.19677024 (ref. 52). Administrative boundary data used for map base layers were obtained from Resource and Environmental Science Data Platform, IGSNRR, CAS and are freely accessible at https://www.resdc.cn/data.aspx?DATAID=205. Population distribution data used for maps were sourced from LandScan and are available at https://doi.org/10.48690/1531770. All third-party datasets are distributed under their respective open-access licenses. Processed outputs and supplementary materials generated during this study are available from the corresponding author upon reasonable request. Source data are provided with this paper.

Code availability

The source code for the main program is available via GitHub at https://github.com/junyuyang7/Composting_strategy_optimization. Executable examples are hosted on the cloud-based computational platform Code Ocean and are available upon request to facilitate peer review and reproducibility.

References

1. *FAOSTAT* (FAO, 2024) https://www.fao.org/faostat/en/#data
2. Mingolla, S. & Rosa, L. Low-carbon ammonia production is essential for resilient and sustainable agriculture. *Nat. Food* **6**, 610–621 (2025).
3. Zhang, X. et al. Nitrogen management during decarbonization. *Nat. Rev. Earth Environ.* **5**, 717–731 (2024).
4. Zhao, S. et al. A precision compost strategy aligning composts and application methods with target crops and growth environments can increase global food production. *Nat. Food* **3**, 741–752 (2022).
5. Chia, W. Y. et al. Sustainable utilization of biowaste compost for renewable energy and soil amendments. *Environ. Pollut.* **267**, 115662 (2020).
6. Bernal, M. P. et al. Current approaches and future trends in compost quality criteria for agronomic, environmental, and human health benefits. *Adv. Agron.* **144**, 143–233 (2017).
7. Zhao, K. et al. Optimizing the management of aerobic composting for antibiotic resistance genes elimination: a review of future strategy for livestock manure resource utilization. *J. Environ. Manage.* **370**, 122766 (2024).
8. Pardo, G., Moral, R., Aguilera, E. & del Prado, A. Gaseous emissions from management of solid waste: a systematic review. *Glob. Change Biol.* **21**, 1313–1327 (2014).
9. Liu, Y. et al. A global meta-analysis of greenhouse gas emissions and carbon and nitrogen losses during livestock manure composting: influencing factors and mitigation strategies. *Sci. Total Environ.* **885**, 163900 (2023).
10. Springmann, M. et al. Options for keeping the food system within environmental limits. *Nature* **562**, 519–525 (2018).
11. Schulte-Uebbing, L. F., Beusen, A. H. W., Bouwman, A. F. & de Vries, W. From planetary to regional boundaries for agricultural nitrogen pollution. *Nature* **610**, 507–512 (2022).
12. Walling, E., Trémier, A. & Vaneeckhaute, C. A review of mathematical models for composting. *Waste Manag.* **113**, 379–394 (2020).
13. Zhao, S. et al. Towards the circular nitrogen economy – a global meta-analysis of composting technologies reveals much potential for mitigating nitrogen losses. *Sci. Total Environ.* **704**, 135401 (2020).
14. Cao, Y. et al. Ammonium enrichment in livestock manure driven by ammonia-abatement practices can reduce nitrous oxide emissions. *One Earth* **7**, 687–696 (2024).
15. Awasthi, M. K., et al. Mitigation of global warming potential for cleaner composting. *Biosynth. Technol. Environ. Chall.* 271–305 (2018).
16. Wang, B., Zhang, P., Qi, X., Li, G. & Zhang, J. Predicting ammonia emissions and global warming potential in composting by machine learning. *Bioresour. Technol.* **411**, 131335 (2024).

17. Huang, N., Ma, S., Zhao, Z. & Zhou, C. Predicting greenhouse gases emissions from decentralized composting by applying explainable machine learning method. *Bioresour. Technol.* **442**, 133726 (2026).
18. He, J. et al. Multi-model machine learning framework for prediction of greenhouse gas emissions during composting of organic solid waste. *Bioresour. Technol.* **443**, 133901 (2026).
19. Liu, L. et al. Knowledge-guided machine learning can improve carbon cycle quantification in agroecosystems. *Nat. Commun.* **15**, 357 (2024).
20. Cao, Y. et al. Mitigation of ammonia, nitrous oxide and methane emissions during solid waste composting with different additives: a meta-analysis. *J. Clean. Prod.* **235**, 626–635 (2019).
21. Mira-Urios, M. Á et al. Composting of olive mill wastewater sludge using a combination of multiple strategies: assessment of improvement in biodegradability, GHG emissions, and characteristics of the end product. *Agronomy* **15**, 808 (2025).
22. Xu, Z. et al. Microbial sources and sinks of nitrous oxide during organic waste composting. *Environ. Sci. Technol.* **58**, 7367–7379 (2024).
23. Wei, Z. et al. Microhabitat drive microbial anabolism to promote carbon sequestration during composting. *Bioresour. Technol.* **346**, 126577 (2022).
24. Ge, J., Huang, G., Li, J. & Han, L. Particle-scale visualization of the evolution of methanogens and methanotrophs and its correlation with CH_4 emissions during manure aerobic composting. *Waste Manag.* **78**, 135–143 (2018).
25. García-Rández, A. et al. Monitoring of greenhouse gas emissions and compost quality during olive mill waste co-composting at industrial scale: the effect of N and C sources. *Waste Manag.* **193**, 33–43 (2025).
26. *Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories* (The Intergovernmental Panel on Climate Change, 2019); https://www.ipcc.ch/report/2019-refinement-to-the-2006-ipcc-guidelines-for-national-greenhouse-gas-inventories/
27. Lee, S.-R. & Kim, G. Assessing ammonia (NH_3) emissions, precursor gas (SO_2, NO_x) concentrations, and source contributions to atmospheric $PM_{2.5}$ from a commercial manure composting facility. *Appl. Sci.* **14**, 11467 (2024).
28. Uwizeye, A. et al. Nitrogen emissions along global livestock supply chains. *Nat. Food* **1**, 437–446 (2020).
29. *FAOSTAT* (FAO, 2020); https://www.fao.org/faostat/en/#data
30. Li, B. et al. Responses of soil organic carbon stock to animal manure application: a new global synthesis integrating the impacts of agricultural managements and environmental conditions. *Glob. Change Biol.* **27**, 5356–5367 (2021).
31. Cao, X. et al. Municipal solid waste compost: global trends and biogeochemical cycling. *Soil Environ. Health* **1**, 100038 (2023).
32. *Waste Generation & Treatment in Ireland in 2018* (Environmental Protection Agency, 2018); https://www.epa.ie/publications/monitoring–assessment/waste/national-waste-statistics/Waste-Statistics-2018-Infographic.pdf
33. Hoy, Z. X., Woon, K. S., Chin, W. C., Van Fan, Y. & Yoo, S. J. Curbing global solid waste emissions toward net-zero warming futures. *Science* **382**, 797–800 (2023).
34. *Global Methane Tracker 2022* (International Energy Agency, 2022); https://www.iea.org/reports/global-methane-tracker-2022
35. Hamrani, A., Akbarzadeh, A. & Madramootoo, C. A. Machine learning for predicting greenhouse gas emissions from agricultural soils. *Sci. Total Environ.* **741**, 140338 (2020).
36. Ji, Z., Zhang, L., Liu, Y., Li, X. & Li, Z. Evaluation of composting parameters, technologies and maturity indexes for aerobic manure composting: a meta-analysis. *Sci. Total Environ.* **886**, 163929 (2023).
37. Samariya, D. & Thakkar, A. A comprehensive survey of anomaly detection algorithms. *Ann. Data Sci.* **10**, 829–850 (2023).
38. Phiwhorm, K., Saikaew, C., Leung, C. K., Polpinit, P. & Saikaew, K. R. Adaptive multiple imputations of missing values using the class center. *J. Big Data* **9**, 52 (2022).
39. Emmanuel, T. et al. A survey on missing data in machine learning. *J. Big Data* **8**, 140 (2021).
40. Kumar, K. & Pande, B. P. Air pollution prediction with machine learning: a case study of Indian cities. *Int. J. Environ. Sci. Technol.* **20**, 5333–5348 (2022).
41. Fu, W. et al. Machine learning-driven prediction of phosphorus removal performance of metal-modified biochar and optimization of preparation processes considering water quality management objectives. *Bioresour. Technol.* **403**, 130861 (2024).
42. Martini, G. et al. Machine learning can guide food security efforts when primary data are not available. *Nat. Food* **3**, 716–728 (2022).
43. Bai, B. et al. Prediction models for bioavailability of Cu and Zn during composting: insights into machine learning. *J. Hazard. Mater.* **471**, 134392 (2024).
44. Shi, S. et al. Improving prediction of N_2O emissions during composting using model-agnostic meta-learning. *Sci. Total Environ.* **922**, 171357 (2024).
45. Deng, P. et al. Development potential of nanoenabled agriculture projected using machine learning. *Proc. Natl Acad. Sci. USA* **120**, e2301885120 (2023).
46. Xiao, L. et al. Spatiotemporal co-optimization of agricultural management practices towards climate-smart crop production. *Nat. Food* **5**, 59–71 (2024).
47. Huijbregts, M. A. J. et al. ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level. *Int. J. Life Cycle Assess.* **22**, 138–147 (2017).
48. Liu, Z., Wang, X., Li, S., Bai, Z. & Ma, L. Advanced composting technologies promotes environmental benefits and eco-efficiency: a life cycle assessment. *Bioresour. Technol.* **346**, 126576 (2022).
49. Silpa, K.et al. *What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050* (World Bank Group, 2018); http://documents.worldbank.org/curated/en/697271544470229584
50. Powell, J. T., Chertow, M. R. & Esty, D. C. Where is global waste management heading? An analysis of solid waste sector commitments from nationally-determined contributions. *Waste Manag.* **80**, 137–143 (2018).
51. Gómez-Sanabria, A., Kiesewetter, G., Klimont, Z., Schoepp, W. & Haberl, H. Potential for future reductions of global GHG and air pollutants from circular waste management systems. *Nat. Commun.* **13**, 106 (2022).
52. Zhang, L., et al. Composting strategy optimization dataset. *Zenodo* https://doi.org/10.5281/zenodo.19677024 (2026).
53. Lebakula, V., et al. LandScan silver edition [data set]. Oak Ridge National Laboratory https://doi.org/10.48690/1531770 (2024).
54. Ba, S., Qu, Q., Zhang, K. & Groot, J. C. Meta-analysis of greenhouse gas and ammonia emissions from dairy manure composting. *Biosyst. Eng.* **193**, 126–137 (2020).
55. Pardo, G., Moral, R., Aguilera, E. & Del Prado, A. Gaseous emissions from management of solid waste: a systematic review. *Glob. Change Biol.* **21**, 1313–1327 (2015).
56. Wang, Y. et al. Mitigating greenhouse gas and ammonia emissions from swine manure management: a system analysis. *Environ. Sci. Technol.* **51**, 4503–4511 (2017).
57. Wang, Y. et al. Mitigating greenhouse gas and ammonia emissions from beef cattle feedlot production: a system meta-analysis. *Environ. Sci. Technol.* **52**, 11232–11242 (2018).
58. Wang, Y. et al. Mitigating ammonia emissions from typical broiler and layer manure management–A system analysis. *Waste Manage.* **93**, 23–33 (2019).

59. Ye, P. et al. Insights into carbon loss reduction during aerobic composting of organic solid waste: A meta-analysis and comprehensive literature review. *Sci. Total Environ.* **862**, 160787 (2023).
60. Zhang, Z. et al. Mitigation of carbon and nitrogen losses during pig manure composting: A meta-analysis. *Sci. Total Environ.* **783**, 147103 (2021).

Acknowledgements
This work was supported by the National Key R&D Program of China (2023YFD1702000), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA28030302), the National Natural Science Foundation of China (32472845, W2412125), the Youth Innovation Promotion Association CAS (2021095), the Hebei Agriculture Research System (HBCT2024230202, HBCT2024270203) and the Natural Science Foundation of Hebei (C2025503005, D2022503014).

Author contributions
Conceptualization, L.Z. and X.W.; formal analysis, L.Z., J.Y., J.L. and X.W.; investigation, L.Z., J.Y., J.L., Q.Z. and Y.R.; methodology, L.Z., J.Y. and J.L.; visualization, L.Z., J.Y. and J.L.; writing—original draft, L.Z. and X.W.; writing—review and editing, L.Z., J.Y., J.L. and X.W.; supervision, X.W., H.Z., Z.B. and L.M.; funding acquisition, X.W., H.Z. and L.M.

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/s43016-026-01361-w.

Correspondence and requests for materials should be addressed to Xuan Wang, Haodi Zhang or Lin Ma.

Peer review information *Nature Food* thanks Jeda Palmer and Susanne Schmidt 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.

nature portfolio

Corresponding author(s): Lin Ma

Last updated by author(s): Apr 21, 2026

Reporting Summary

Nature Portfolio 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 Portfolio 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	To extract specific data from figures in the selected literature, GetData Graph Digitizer 2.22 was employed.
Data analysis	Data processing and visualization were performed using a combination of Python 3.8, Microsoft Excel 2024, OriginLab 2026, and ArcGIS 10.6. The core algorithms and code developed for this study are publicly available in a GitHub repository (https://github.com/junyuyang7/ Composting_strategy_optimization).

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 Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The datasets supporting the findings of this study are publicly accessible. The core dataset required to execute the main program has been permanently archived in Zenodo and is available via the persistent identifier(https://doi.org/10.5281/zenodo.19677024).Administrative boundary data used for map base layers were obtained from Resource and Environmental Science Data Platform, IGSNRR, CAS and are freely accessible at https://www.resdc.cn/data.aspx?DATAID=205. Population distribution data used for mapwere sourced from LandScan and can be downloaded via https://doi.org/10.48690/1531770. All third-party datasets are distributed under their respective open-access licenses. Processed outputs and supplementary materials generated during this study are available from the corresponding author upon reasonable request.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	not applicable
Reporting on race, ethnicity, or other socially relevant groupings	not applicable
Population characteristics	not applicable
Recruitment	not applicable
Ethics oversight	not applicable

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

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

Ecological, evolutionary & environmental sciences study design

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

Study description	This study employs machine learning to analyze 848 global composting experiments, this study identified 19 key management parameters driving emissions of ammonia (NH3), nitrous oxide (N2O), methane (CH4), and carbon dioxide (CO2). We estimated annual global losses at 747 kt NH3-N, 81 kt N2O-N, and 592 kt CH4-C from composting, these non-CO2 emissions were equivalent to approximately 61 Mt CO2eq per year. By optimizing strategies—such as adjusting aeration, additives, and C/N ratios—the composting chain could transition from a net emitter (40.1 Mt CO2eq) to a C sink (-15.1 Mt CO2eq), while conserving nutrients for crops. These innovations directly advance sustainable agriculture through waste-to-resource systems that boost food security and climate resilience.
Research sample	In this study, the modeling data include 848 observations reported in 171 peer-reviewed publications from 1993 to 2023, and FAOSTAT data and multiple national data sources were integrated to assess the emission reduction potential of optimized management strategies for organic waste composting within the global food system.
Sampling strategy	In this study, the experimental data selected from the literature as the data source must meet the following criteria: 1) the composting process must be complete; 2) laboratory-scale incubation experiments are not included; and 3) the study must include emission factors for at least one of NH3, N2O, CO2 and CH4. For the selection of optimization strategies for composting, five screening approaches were considered, including minimizing global GHGs emissions, Nr mitigation, and minimizing the emission of individual gases.
Data collection	The data used in this study were sourced from global and regional statistics, surveys, reports, published research papers containing field experiment results, and model databases.
Timing and spatial scale	This study conducted model training based on a dataset compiled from available literature in the field of composting from 1993 to 2023. Additionally, using 2020 as the baseline, it performed estimations and analyses on organic waste composting within the global food system.
Data exclusions	No data were excluded from the analyses.
Reproducibility	This study is not based on field experiments, but rather on a retrospective analysis of published observational composting data and secondary datasets (e.g., FAOSTAT, national statistics, and model databases). Therefore, no new experimental procedures were conducted that would require classical replication through repeated trials. However, to ensure computational reproducibility, all 848 data points were traceable to 171 peer-reviewed publications, enabling independent verification. We also conducted 200 random runs for each emission factor using bootstrap and stratified sampling methods, and with 1000 runs Monte Carlo simulation confirmed result stability. Independent re-runs of the full pipeline consistently reproduced the findings, with only minor, expected

	stochastic variation. Thus, although no physical experiments were conducted, all analytical results were successfully and reliably replicated.
Randomization	This study employs machine learning to analyze observational composting experiments and secondary data drawn from global statistics, published literature, and model databases. It does not involve biological samples, living organisms, or human participants, nor does it entail experimental interventions or group assignments. Consequently, the notion of allocating subjects into treatment or control groups is not applicable in this study.
Blinding	Blinding was deemed unnecessary, as the data utilized in this study was neither subjective nor susceptible to researcher biases.

Did the study involve field work? ☐ Yes ☒ No

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
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

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

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*