Arable soil nitrogen dynamics reflect organic inputs via the extended composite phenotype, not the metagenome associated with nitrogen transformations.

Andrew L. Neal[1†], Harry A. Barrat[1], Aurélie Bacq-Lebreuil[2‡], Yuwei Qin[3], Xiaoxian Zhang[4], Taro Takahashi[1,5], Valentina Rubio[6,7], David Hughes[8], Ian M. Clark[4], Laura M. Cárdenas[1], Laura-Jayne Gardiner[9], Ritesh Krishna[9], Margaret L. Glendining[8], Karl Ritz[2], Sacha J. Mooney[2], John W. Crawford[10].

[1]Net Zero and Resilient Farming, Rothamsted Research, North Wyke, EX20 2SB, UK;

[2]School of Biosciences, The University of Nottingham, Sutton Bonington, LE12 5RD, UK;

[3]Department of Environmental Sciences, Wageningen University, 6700 HB Wageningen, The Netherlands;

[4]Sustainable Soils and Crops, Rothamsted Research, Harpenden, AL5 2JQ, UK;

[5]Bristol Veterinary School, University of Bristol, Langford, BS40 5DU, UK;

[6]Programa de Producción y Sustentabilidad Ambiental, Instituto Nacional de Investigación Agropecuaria (INIA), Estación Experimental INIA La Estanzuela, Colonia, Uruguay;

[7]School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA;

[8]Intelligent Data Ecosystems, Rothamsted Research, Harpenden, AL5 2JQ, UK;

[9]IBM Research Europe - Daresbury, The Hartree Centre, Warrington, WA4 4AD, UK;

[10]Adam Smith Business School, University of Glasgow, Glasgow, G12 8QQ, UK.

[†]andy.neal@rothamsted.ac.uk; [‡]present address: Genesis, 4 rue de l'église, 27440, Lisors, France.

Abstract

Achieving food security whilst minimising agriculture's influence on climate and biodiversity requires resilient systems, predicated on improved nutrient-use efficiency, water and nutrient storage in soils and reducing gaseous emissions. Success requires sufficient understanding of coupled nitrogen and carbon metabolism in soils, associated influences on soil structure, and of processes controlling nitrogen-transformations at scales relevant to microbial activity. We show the influence of organic matter on arable soil nitrogen transformations and balances can be understood by integrating metagenomic data with soil structural parameters. System function can be predicted from theory of soil as an extended composite phenotype, where physical and biotic interactions result in

emergent organisation and function. The approach provides mechanistic explanation of why organic matter is effective in reducing nitrous oxide losses while supporting system resilience. The relationship between organic carbon, soil connected porosity and flow rates at scales relevant to microbes suggests significant increases in efficiency and resilience could be achieved at lower organic carbon stocks than currently envisaged.

Main

Agricultural production must balance ostensibly incompatible demands. Ideally, it must provide sufficient, nutritious food to increasing human populations, reduce its environmental footprint, restore biodiversity, and attenuate greenhouse gas (GHG) emissions and associated climate change. Soil is simultaneously the most important terrestrial sink and source of GHG, particularly the long-lived gases carbon dioxide and nitrous oxide (N_2O), and shorter-lived methane[1]. As N_2O constitutes a loss of an important plant nutrient from the system, nitrogen dynamics in soil are particularly important in linking climate regulation and food security[2]. Improving nutrient use efficiency (NUE) such that GHG emissions resulting from fertiliser production and use falls by 40% *per* unit of food produced is predicted to be more effective at reducing emissions from the food system than higher-yielding crops or a 50% reduction in food waste[3]. NUE and carbon sequestration, are closely coupled in soils[4], co-regulated by microbial metabolism which, in turn, is governed by oxygen availability and co-location of nitrogen in, and together with, organic matter. The distribution of oxygen, in turn, is determined by the physical soil pore structure , its moisture content, and the distribution of potential microbial respiration[5]. Therefore, an understanding of how management impacts on the partitioning of carbon and nitrogen between storage in soil and emissions requires and understanding of the dynamics of both the physical and biotic state. The '*extended composite phenotype*' theory of soil provides the required framework.

Important soil functions, particularly transport capacity, metabolic efficiency, and resource supply resilience to plants, acting across a broad range of scales, are dependent upon the connectivity and heterogeneity of pore space, emphasizing the importance of a pore-centric view of soil architecture[6].

56 The theory of soil as an extended composite phenotype is based on detailed observations and
57 modelling of the interactions between structural genesis and microbial community structure and
58 function[7]. Soils exhibit spontaneous emergence of multi-scale self-organisation driven by endogenous
59 feedback between pore space architecture and microbial metagenetic—rather than taxonomic—
60 states[7]. Microbiological activity influences soil pore architecture at scales below approximately 80
61 μm. Since these are the scales that affect the relative balance and distribution of air and water in soil,
62 these changes also influence the nature of microbial activity, especially the relative activity of aerobic
63 and anaerobic metabolism. The resulting process-form state is a consequence of the self-organisation
64 of the integrated biophysical system of soil[7,8]. The emergent state is affected by incorporation of OM
65 into soil because this drives a tightly coupled feedback involving changes to the process-form state
66 associated with altered gene assemblages and soil texture – hence the extended phenotype[7-10] of the
67 soil system is an irreducible composite reflecting physical and biotic feedbacks.
68 Based upon our previous observations, the soil extended composite phenotype incorporates microbial
69 gene assemblages, soil structural parameters, hierarchical flux processes of gases and water, and
70 overall system function including the likelihood of anoxia. In soil with a high throughput of OM,
71 process-form states are characterised as having a high porosity and a high degree of pore
72 connectivity[11,12]. Reduced levels of OM are associated with less extensively developed soil pore
73 networks. The resultant altered state modifies diffusivity and O_2 permeability[7,13-16]. This pore-centric
74 perspective is essential when considering the consequences of incorporating OM into arable soils for
75 NUE and GHG emissions. The extended composite phenotype of soils containing low stocks of
76 organic carbon is characterised by greater proportions of anoxic pore space—the principal control of
77 biological denitrification activity[17]—and increased abundance of genes in the metagenome associated
78 with dissimilatory respiration of nitrate and nitrite[7]. These factors effectively link the fates of nitrogen
79 and carbon in soil.
80 Here, we provide a further test of the theory of soil as an extended composite phenotype, to provide a
81 mechanistic explanation of the links between carbon and nitrogen metabolism and show how both
82 physical and genetic factors must be accounted for. According to this theory, we hypothesise that
83 arable soils receiving high organic inputs would have a highly connected process-form state resulting

in reduced losses of stored nitrogen as N_2O through anaerobic metabolism. Conversely, soils receiving low organic inputs would result in a poorly connected state, reduced oxygen flux and higher gaseous losses of nitrogen. Furthermore, using this framework, we seek to explain the link between carbon dynamics and non-equilibrium NUE, calculated as the balance of nitrogen inputs, off-takes and accumulation in soil. We hypothesise that offsetting losses of N_2O from soil would be associated with higher stocks of soil nitrogen, along with carbon and water storage—key factors in conveying resilience in rainfall-limited production systems. To test our hypotheses, we studied arable soils subjected to consistent management for over 160 years, comparing soils which had received a range of continuous organic or inorganic fertilisation over this period.

Results and discussion

Emergent soil process-form states

Co-metabolism of carbon and nitrogen – Since experiment establishment (1843), soils subject to different management or fertilisation (Table I) have developed distinct soil organic carbon (SOC) and total nitrogen (N_{tot}) stocks (Supplementary Fig. 1). Woodland and proximal grassland soils, together with arable soil receiving composted farmyard manure (FYM) were the only soils to display net SOC and N_{tot} accumulation (Fig. 1). Grassland and FYM-amended arable soils both contained over 70 Mg ha^{-1} SOC and over 6 Mg ha^{-1} N_{tot}, significantly more than other treatments. There was a positive geometric mean functional relationship between SOC and N_{tot} (Fig. 2), corresponding to a C:N ratio of 11.4. Despite widely varying quantities and qualities of carbon and nitrogen inputs, this indicates coupling via the same metabolic pathways and mechanisms of storage in soil. Metabolic and storage constraints in soil are both linked to microenvironmental conditions, *via* a strong association between SOC inputs and soil process-form relationships[7]. Therefore, we quantified micrometre-scale structure of a subset of soils to understand the biophysical feedbacks in the co-metabolism of carbon and nitrogen.

Links between nitrogen storage and physical processes – Parameters relating to soil architecture were determined in grassland, woodland, FYM, and inorganically fertilised arable soils

receiving 144 kg-N ha^{-1} yr^{-1} as ammonium nitrate, phosphorus and potassium (^{144}NPK), 192 kg-N ha^{-1} yr^{-1} and potassium but no phosphorus (^{192}NK), and soil receiving no nitrogen fertilisation but phosphorus and potassium (PK). Significant treatment-dependant differences were observed for parameters generated directly from X-ray computed tomography of pore networks in each soil (total [P_t] and connected [P_c] porosity) and those derived from simulation of the permeability (k), normalized effective oxygen diffusion coefficient (D_e'), and hydraulic conductivity (K) within the pore networks (Table II). Collectively, these measures describe the dynamical state of soil pore space and the maximum potential rate at which resources can move through the networks, *i.e.*, the capacity for flux. Grassland, woodland and FYM-amended arable soils were typified as having more extensive and more connected pore networks than inorganically fertilised arable soils. We observed a power-law relationship between P_c and K: increased SOC was associated with increases in both parameters (Fig. 2B). Regions of this relationship correspond to the process-form states of the various soils. This infers that—in the case of the soils studied here—for P_c between 0.05 and 0.4, relatively small changes in geometry, characterised by P_c, result in substantial increases in K, characterising flow rate. This power law therefore has profound implications for soil management strategies.

The effect of soil management was most evident on permeability (k) based upon treatment effect size (ω^2, Table II). We simulated the anoxic proportion of each soil across a range of matric potentials (ψ_m). Addition of FYM to arable soils resulted in a matric potential–anoxic space profile distinct from inorganically fertilised arable soils (Fig. 2C). Inorganically fertilised ^{192}NK and ^{144}NPK soils, as well as soil that had received no fertilisation (referred to as *nil*, Table I) all presented large proportions of anoxic space under relatively dry conditions (ψ_m 50–65 kPa) and were predicted to be completely anoxic at ψ_m between 38.4-32.0 kPa. In contrast, the process-form state of FYM-amended soil exhibited aspects of woodland and grassland soils. Under relatively dry conditions FYM soil had low proportions of anoxic space more typical of grassland soil. At increased moisture content, both soils were completely anoxic at 22.6 kPa; woodland soil being so at 20.2 kPa.

As with an affiliated ley-arable experiment[7], organic-carbon rich soil developed a distinct process-form state, typified by a greater proportion of connected pores, and increased hydraulic conductivity (Fig. 2B). Modelling predicts this to be a more oxygenated pore network (Fig. 2C) by

virtue of its greater capacity to transport O_2. Annual addition of FYM (35 Mg ha^{-1}) to the soil has resulted in SOC and N_{tot} contents, and process-form states resembling unmanaged woodland and grassland, despite regular physical disturbance by inversion tillage. This combined evidence is consistent with our hypothesis regarding the influence of organic carbon in soils in creating a more highly connected and oxygenated structure.

Consequences of soil processes-form state for potential biological function – Nitrogen transformation-associated metagenomes were determined in grassland, woodland, FYM, ^{144}NPK, ^{192}NK and PK soils from shotgun metagenomic approaches. There was a significant treatment effect upon gene assemblages (PERMANOVA, 99,999 permutations: *pseudo*-$F_{5,12}$ =22.5, p_{perm}=1x10^{-5}). *Post hoc* pairwise comparisons indicated no significant difference in gene assemblages between ^{192}NK and PK soils. All other comparisons were significantly different.

Hierarchical clustering of soils based upon centred log-ratio transformed counts of read numbers matching each identified gene (*i.e.*, relative abundance) demonstrated gene assemblages of unmanaged woodland and grassland soils were distinct from arable soils (Fig. 3). Genes separated into two broad clusters based upon their distribution between unmanaged and arable soils. The first included genes associated with amino-acid metabolism and other nitrogen-assimilation pathways, including nitrate assimilation (KEGG module M00615) and assimilatory nitrate reduction (M00531). Several genes associated with dissimilatory nitrate reduction (M00530) and denitrification (M00529) were also associated with this cluster. Relative abundance of these genes was greater in woodland and grassland soils—which were most alike—than arable soils (Fig. 3). More subtly, genes associated with nitrate assimilation were relatively more abundant in FYM-amended than inorganically fertilised arable soils.

The second gene cluster was generally more abundant in arable soils. This comprised genes associated with several modules, including nitrification (M00528), denitrification (M00529), dissimilatory nitrate reduction (M00530) and complete nitrification (M00804), (*c.f.* Supplementary Information). This pattern of limited genotypic differences between arable soils is broadly consistent with known responses of microbial communities to nitrogen fertilisation[18-21]. This is evidence for

direct influence of process-form states upon nitrogen-associated gene assemblages in soil, particularly the association of nitrogen assimilatory pathways with soils having the greatest SOC and N_{tot}, connected porosity, and hydrodynamic conductivity. The greater relative abundance of genes associated with nitrogen assimilation in FYM-amended soil is consistent with this soil having a similar process-form state (Fig. 2B) and oxygen status (Fig. 2C) to woodland and grassland soils. Soils having low SOC (and N_{tot}) and thus low connected porosity and hydrodynamic conductivity were associated more with dissimilatory pathways, utilizing oxidised forms of nitrogen as alternative electron acceptors for respiration. Within arable plots particularly, there was a subtle shift in genes associated with denitrification and dissimilatory nitrate reduction: NO-forming and cytochrome *c*-type nitrate reductases (*nirK* and *napAB* respectively) were relatively more abundant in FYM-amended than inorganically fertilised soils. These genes are typically associated with more oxygenated environments than the functionally equivalent but structurally dissimilar genes *nirS* (*c*-type cytochrome) and the nitrite oxidoreductase *nxrAB*[22-25] and were most abundant in grassland and woodland soils respectively. In addition, FYM had the highest relative abundance of genes of the high-affinity nitrate-binding transporter (*nrtABC*) of all arable soils. These genes were otherwise most abundant in grassland soils. Nitrite reductases typical of more reducing environments (*nirS*, *nxrAB*) were indicative of arable soils in general, particularly ^{144}NPK, ^{192}NK treatments. In FYM-amended soil, the more oxygen-rich process-form state (Fig. 2C) influences the gene assemblage such that it shares similarities to those of grassland and woodland soils.

Assessment of gene abundance associated with nitrogen transformations by quantitative PCR provides evidence of the pattern and rate of nitrogen cycling processes in soil[26]. However, the relative abundance of different genes discussed above (and Supplementary Information) generated from metagenomics cannot in themselves be used as the basis for estimates of process rates[27]. We assumed relative gene abundance reflects the *likelihood of a process* with which they are associated when environmental conditions allow. To test this assumption, we measured N_2O emissions in the field, relating them to relative gene abundance and predictions of anoxic pore space (Fig. 2c).

Nitrous oxide emissions from soil – Since lower levels of organic carbon are associated with less extensively-developed soil pore networks[11,12] (Table II) and the resultant low diffusivity restricts oxygen permeability[7,13-16] (Fig. 2C), elevated N_2O emissions are often associated with poor soil architecture. We placed static chambers[28] in the field on FYM, ^{240}NPK, ^{192}NK, ^{144}NPK and PK soils between April and November 2019 to compare soil N_2O emissions. For each measurement date, greatest emissions were typically measured from ^{240}NPK soils and least from PK soils (Supplementary Fig. 2). We detected no significant influence of soil temperature (ANCOVA; $F_{1,134}$=0.901, p=0.344) or potential soil moisture deficit (ANCOVA; $F_{1,134}$=0.272, p=0.603) upon N_2O emissions. Comparing soils directly, there was a significant influence of arable soil management (ANOVA; $F_{4,116}$=8.7, p=3x10^{-6}) upon soil mean daily N_2O emission (Fig. 4). FYM-amended soil emitted significantly less N_2O than ^{240}NPK soil (t=3.1; p=0.005), despite receiving similar annual nitrogen inputs and having more than double the nitrogen stocks of inorganically fertilised soils (Supplementary Fig. 1).

Non-equilibrium nitrogen use efficiency of the soil-plant system – We hypothesised that differences in extended composite phenotype of FYM-amended and inorganically fertilised arable soils would be reflected in historical nitrogen fluxes and stocks through the soil-plant systems. Traditionally, NUE is expressed as a ratio between the amount of nitrogen introduced to and harvested from a system within a single season[29-31]. This implicitly assumes a system is operating under a long-term equilibrium, where soil nitrogen stocks remain temporally constant. For the inorganically fertilised plots this condition is largely satisfied, due to the long trial history meaning annual production of soil organic nitrogen virtually equals annual nitrogen mineralisation[32,33]. However, when this condition is not satisfied—FYM-amended soils continue to accumulate nitrogen (Fig. 1B)—applied nitrogen stored in soil beyond a single cropping cycle is incorrectly considered ‘wasted’. Although nitrogen which accumulates in soil is not recovered by the immediate crop, it may contribute to system resilience by supporting production of subsequent crops. To account for this, we calculated non-equilibrium NUE for treatments, apportioning nitrogen introduced to each system between that taken up by the crop, stored in soil and ‘lost’ from the soil-plant system. For lost nitrogen, gaseous losses resulting from

denitrification may be up to double those via leaching[34]. Since ^{240}NPK and FYM receive similar levels of nitrogen inputs (but differ in levels of readily available nitrogen), the effect of organic carbon input can be evaluated. Mean annual nitrogen fluxes through the three pools following addition to the different winter wheat crop systems (2000-2015) demonstrate progressively greater losses of nitrogen inputs with increasing fertilisation (Fig. 5A). Whilst total nitrogen associated with wheat grain and straw increases as nitrogen inputs increase, the proportion of total assimilated nitrogen reduces (Supplementary Fig. 3), indicating progressively reduced long-term NUE. Nitrogen inputs required to generate a 1 Mg grain harvest show similar trends (Fig. 5B). Set alongside inorganically fertilised systems, and particularly ^{240}NPK, the FYM system shows altered allocation of added nitrogen to each pool. Reduced total (Fig. 5A) and proportional (Supplementary Fig. 3) allocation of nitrogen to lost pools and increased in nitrogen stocks (*cf*. Figs. 1 and 5A), suggest greater NUE. This increased efficiency is tempered by reduced nitrogen allocation to crops and grain yields (Fig. 5). The 2000-2015 average grain yield from FYM-amended soil was 1.1 Mg ha^{-1} less than from ^{240}NPK fertilised soil (Q=4.2; p=0.035). This is consistent with lower area-scaled (but greater yield-scaled) N_2O emissions from soils receiving organic amendments rather than inorganic fertiliser[35]. Measured and simulated micropore-scale behaviour of soil extended composite phenotypes reflect our field-scale description of nonequilibrium NUE, particularly reductions in the absolute and proportional size of the lost nitrogen pool. Our data suggest that the application of FYM to soil is directly related to the reduction of nitrogen losses since we observe a 1.5-fold increase in soil P_t and a four-fold increase in k in FYM-amended over inorganically fertilised soils.

The extended composite phenotype and system resilience - Consistent with our hypothesis that nitrogen balance in soils would be altered under different emergent process-form states, we observe that the reduced area-scaled N_2O emissions in FYM amended soils associated with high connected porosity and predicted greater oxygen availability are associated with greater N_{tot} stocks over the long term. By combining physical structure with metagenetic and phenotypic description, we invoke a more fundamental and mechanistic conceptualisation of the role of organic matter in soil function. We demonstrate that the nitrogen-associated metagenome responds to nitrogen fertilization markedly, but

that form of nitrogen input exerts relatively little influence upon redox sensitive *nirK-napAB* and *nirS-nxrAB* gene abundances. Given the magnitude of the changes we observe in process-form state, it is likely that OM controls N_2O emissions by influencing gene expression (phenotype) rather than altering gene abundance (genotype). In effect, reduced nitrous oxide losses are a co-benefit of increased organic inputs. This regulation results in nitrogen accumulation in soil conferring a degree of resilience, consistent with previously observed reductions in the need for inorganic nitrogen application to achieve a target yield in organic carbon-rich soils[36]. Evidence for higher SOC stocks being associated with higher crop yields is equivocal[37-38]. However, resilience of rainfed crop yields to drought does appear to be related to higher SOC stocks[39,40], although in the soils studied here such effects are only observed where nitrogen availability limits crop yield[41].

Conclusions -- Soil presents a multivariate conundrum in which simple, linear cause-effect processes—however conceptually appealing—are difficult to reconcile with field data. Our adoption of a novel non-linear and systems-level approach, founded on quantification of the process-form state and interactions with the metagenome, shows that the behaviour of the soil system cannot be understood by studying the behaviour of the physical and biotic components in isolation. Our emphasis is on a parsimonious explanation for our observations. It shows that the form of nitrogen fertilisation in arable systems does little to influence nitrogen-metabolism-associated gene assemblages, but control over expression is exerted by the emergent process-form state which is dependent upon organic matter inputs. Combined evidence suggests a mechanistic basis for why regular addition of organic matter to arable soils is effective in reducing losses of nitrogen as N_2O to the atmosphere. Whilst the applications of FYM used in our experimental system are impractical[42], the power-law relationship observed in Fig. 2 implies that similar levels of efficiency and resilience could be achieved at lower SOC input and stocks than used here. One consequence of close coupling of carbon and nitrogen metabolism is that increased N_2O emissions may offset any benefits for limiting global warming accruing from actively increasing SOC stocks[43,44]. Our data suggest that optimally efficient systems can achieve both increased SOC stocks and reduced N_2O emissions. The extended composite phenotype concept provides an

integrative explanation why this might be so, based upon a biological and pore-centric understanding of soil processes.

Methods

Field experiment and soil sampling – Soil used to generate physical data relating to soil structure and biological data derived from metagenomics was sampled from arable soils which had received inorganic fertiliser inputs ranging from no inputs to different combinations of inorganic nitrogen (ammonium nitrate), phosphorus and potassium, comparing rates of nitrogen inputs from 0 to 240 kg ha^{-1} yr^{-1} from contrasting treatments of the Broadbalk Long-Term field experiment established in 1843 (51˚48'35" N, 00˚22'30" W). The experiment tests the long-term consequences of different fertiliser and manure applications on the yield of winter wheat, as well as the effect of cessation of all cultivation from a part of the site in 1882. Detailed description of the treatments compared in this study are provided in the Supplementary Information: comprehensive details regarding the experiment history are provided by Macdonald *et al.*[45]. We also studied two soils which had not been subject to management for over a century: part of the original Broadbalk experiment which since 1882 has been allowed to revert naturally to a woodland (referred to as the woodland treatment), and plots of the Highfield Ley-Arable experiment managed as a mown sward[45] (referred to as the grassland treatment). Details of the management associated with the different soils compared in his work are provided in Table I. Treatments are not replicated across the Broadbalk experiment. Therefore, samples associated with metagenomics and X-ray computed tomography (both described below) are by necessity *pseudo*-replicates. Assessment of differences between treatments as a consequence are relative to the pooled within-plot variability in both ANOVA and PERMANOVA tests, rather than some form of between-plot variability, and we cannot statistically separate spatial from "treatment" effects. Our approach therefore tests the hypothesis that there are differences in measured and simulated soil parameters, and assemblages and relative abundance of nitrogen cycle-associated genes in the different plots at the Broadbalk site. We cannot conclude, based on statistical inference, that any differences observed result from different fertility management directly. However, it is reasonable to interpret our observations of between-plot differences in soil structure and relative gene abundance

within the context of the broader data relating to SOC and N_{tot} stocks, held by the *e*-RA electronic database, and N_2O emissions—all of which are temporally replicated—to suggest explanations for the patterns we observe.

Soil structure and hydrodynamic behaviour

X-ray computed tomography and image analysis – We generated X-ray CT images from grassland, woodland, FYM, ^{144}NPK, ^{192}NK and nil treatments at 1.5 µm resolution and scales relevant to microbes (10^0-10^2 µm), requiring imaging of 0.7 – 2.0 mm diameter soil aggregates. Aggregates (n = 9) were selected at random from soil collected from each plot. Each was scanned using a Phoenix Nanotom system (GE Measurement and Control solution, Wunstorf, Germany) operated at 90 kV, a current of 65 µA and at a voxel resolution of 1.5 µm. Initial image analysis was performed using Image-J. Images were threshold-adjusted using a bi-level bin approach[47] using QuantIm version 4.01. Porosity was calculated directly from threshold-adjusted binary images, pore connectivity was determined according to Vogel *et al.*[47].

Pore network permeability, hydraulic conductivity and oxygen diffusion in soil – The permeability (k) and hydraulic conductivity (K) of soil pore networks, and the effective diffusion coefficient (D_e) for oxygen within the pores was calculated for each soil structure imaged by X-ray CT using lattice Boltzmann simulation[47,48]. Permeability and hydraulic conductivity were calculated as described by Zhang *et al.*[15]. For oxygen diffusion, hierarchical soil structures revealed by X-ray CT images indicate that gaseous oxygen in the atmosphere moves into soil primarily through its inter-aggregate pores and is then dissolved in water prior to moving into aggregates, largely by molecular diffusion. Since gaseous oxygen diffuses up to 10^3-fold more quickly than oxygen dissolved in water, microbial community activity is constrained mainly by oxygen diffusion within aggregates. The ability of aggregates to conduct dissolved oxygen and other soluble substrates depends on the intra-aggregate pore geometry, and we quantified it with effective diffusion coefficients calculated directly by mimicking solute movement through the pore geometry using numerical simulations. The movement of solutes, including oxygen, within the pore geometry is assumed to be diffusion dominated. To account for the effect of temperature upon diffusion, we calculated a normalized

effective diffusion coefficient D_e' dividing the effective diffusion coefficient by the molecular diffusion coefficient of oxygen, D, in water at the same temperature, *i.e.*, $D_e' = D_e/D$[15,16]. Detailed description of pore-scale simulation is provided in the Supplementary Information.

Modelling anoxia within soils – Having established D_e' for oxygen in the soil pore networks of each soil, we simulated oxygen consumption by microbes under various levels of water saturation[7]. The first step of simulation was to determine water distributions in the pore networks under different matric potentials (ψ_m). Once water distributions were determined for a given ψ_m, we simulated oxygen dissolution at air-water interfaces and then diffusion towards solid-water interfaces where it was reduced by microbial respiration. Microbial consumption was assumed to occur in water-filled voxels adjacent to the water-solid wall and described by a Monod kinetic equation. Oxygen diffusion and reduction was simulated to steady state. As the development of anaerobic volume is a balance between oxygen diffusion within the pore network and microbial consumption, we simulated two scenarios: fast ($k' = 1x10^{-2}$) and slow ($k' = 1x10^{-4}$) microbial consumption of oxygen to determine whether the relative anaerobicity of soils under the same ψ_m was a consequence of their structures and not microbial respiration rate. Once each system had reached a steady state, we considered sites where concentration of dimension-less dissolved oxygen was less than 20% to be anaerobic[49]. We repeated the procedure to achieve different water distributions calculated by varying ψ_m and then calculated the proportional change in the volumetric anaerobic sites with the ψ_m for both the fast and slow microbial reactions. Detailed description of these steps is provided in the Supplementary Information.

DNA extraction, sequencing and quality control – Three *pseudo*-replicate samples of soils from the grassland, woodland, FYM, ^{144}NPK, ^{192}NK and PK treatments were sampled in October 2015. Soil community DNA was extracted from a minimum of 2 g of thawed soil the using MoBio PowerSoil DNA isolation kits (Mo Bio Laboratories, Inc. Carlsbad, CA). 10 µg of high-quality DNA was provided for sequencing for each of the samples. Shotgun metagenomic sequencing of DNA was performed using 150-base paired-end chemistry on an Illumina HiSeq 2500 sequencing platform by Beijing Novogene Bioinformatics Technology Co. Ltd. (Beijing, China). The generated raw

sequences were limited to a minimum quality score of 25 and a minimum read length of 70 bases using Trimmomatic[50].

Bioinformatic analysis of metagenome sequences - To assess general abundance of genes in metagenomes, we mapped individual metagenomic sequences to the RefSeq non-redundant (NR) protein database held at NCBI (downloaded August 22nd, 2018) using DIAMOND version 0.8.27[53] in BLASTX mode imposing a bitscore cut-off of 55. For each sequence, only the match with the highest bitscore was considered. Sequences not matching the NR database were considered currently unclassified. MEGAN Ultimate version 6.10.2[52] was used to associate metagenome sequences with Kyoto Encyclopaedia of Genes and Genomes[53] (KEGG) functional orthologs (ko) and modules (M). From all the reads binned to at least one KEGG orthologous group, we selected those associated with nitrogen metabolism (ko00910), amino acid metabolism (ko09105) and metabolism of other amino acids (ko09106) for detailed study of distribution differences between soils. Nitrous oxide is a product of autotrophic and heterotrophic nitrification pathways as well as denitrification. Emission rates differ markedly, both between processes and at different soil saturations[54]. Emission rates of N_2O due to denitrification are far greater than from either nitrification pathway[54]. Not only is nitrification a minor source of N_2O, but arable systems also exhibit the lowest maximum potential nitrification-derived N_2O emissions across a broad range of ecosystems, and this declines with increasing management intensity[55]. For these reasons, we assumed that N_2O emissions were predominantly driven by the influence of oxygen upon anaerobic respiration *via* the denitrification pathway. For each module in the KEGG ontology, genes may be associated with several sub-modules. These are not formed of exclusive sets of genes and so are likely to reflect a range of function and ecophysiologies. In addition, the ontology does not distinguish between *amoABC* associated with nitrification and *pmoABC* associated with methane oxidation. Details of individual genes included in the analysis is provided in the Supplementary Information and the distribution of genes across different modules are shown in Fig. 3.

Emissions of nitrous oxide from soil – All nitrogen fertilisers were applied on April 12th, 2019, FYM was applied 17th September 2018 and 23rd September 2019. Measurement of gaseous emissions

from ^{240}NPK, FYM and PK soils were taken on eleven dates between April 11th and October 7th, 2019. Measurements from ^{144}NPK and ^{192}NK soils were taken on eight dates between April 23rd and October 7th, 2019. Gas sampling was performed using in-field static chambers[28]. Three chambers (dimensions 40 cm x 40 cm x 25 cm height) were inserted to a depth of 5 cm in soil of each treatment. Details of sampling times and frequency are provided in the Supplementary Information. N_2O concentrations in gas samples collected in the field were analysed on a PerkinElmer Clarus 500 Gas Chromatograph (GC) equipped with a TurboMatrix 110 automated headspace sampler, fitted with an electron capture detector set at 300 °C for N_2O analysis. The static chamber approach used here is prone to errors derived from GC instrumental noise and temperature and pressure changes within the chamber leading to artefactual negative fluxes of N_2O[56]. To avoid these artefacts, only non-zero flux estimates were used to analyse differences in N_2O emissions between treatments: negative fluxes were removed from the analysis. We summed total N_2O emissions and tested for differences in the mean N_2O emissions between treatments for the sampling period assessed on an area basis per day.

Non-equilibrium nitrogen use efficiency of the soil-plant system – The balance of nitrogen inputs, off-takes in wheat grain and straw and accumulation in arable soils was estimated for FYM, ^{240}NPK, ^{192}NPK, ^{144}NPK and PK Broadbalk fertiliser treatments between 2000 and 2015. Historical data relating to nitrogen inputs (as ammonium nitrate fertiliser or FYM, within seed grain and atmospheric deposition), off-takes (as harvested grain and straw), and soil nitrogen stocks were acquired from the *e*-RA managed database of data from Rothamsted's long-term experiments (data available from http://doi.org/10.23637/rbk1-yldS10115-01). The end-of-season fate of annual nitrogen additions in the form of inorganic fertiliser or FYM introduced to the Broadbalk continuous wheat plots was separated into three pools: (1) that taken up for the current season's production (*i.e.* nitrogen content in straw and harvested grain); (2) the pool held within the system for potential use in the future (*i.e.* change in soil nitrogen stock); and (3) that lost from the system without being utilised (*e.g.* as N_2O emitted to the atmosphere or nitrate leached to groundwater). The agronomic performance of each plot was then evaluated by nonequilibrium nitrogen use efficiency, here defined as the ratio between nitrogen input and (1)+(2) above.

Availability of data

Data relating to the Broadbalk and Highfield long-term experiments can be accessed *via* the electronic Rothamsted Archive (http://www.era.rothamsted.ac.uk/experiment/rbk1 and http://www.era.rothamsted.ac.uk/experiment/rrn1, respectively). Historical data relating to nitrogen inputs (as ammonium nitrate fertiliser or FYM, within seed grain and atmospheric deposition), off-takes (as harvested grain and straw), and soil nitrogen stocks are available from http://doi.org/10.23637/rbk1-yldS10115-01. All soil images are available upon reasonable request from the corresponding author. Code relating to lattice Boltzmann simulation is available upon request from Xiaoxian Zhang, Rothamsted Research. Sequence data associated with this research have been deposited in the European Nucleotide Archive https://www.ebi.ac.uk/ena/browser/view/PRJEB43407?show=reads.

Acknowledgements

This research was supported by UK Research and Innovation's (UKRI) Biotechnology and Biological Science Research Council (BBSRC)-funded Soil to Nutrition strategic program (BBS/E/C/000I0310 and BBS/E/C/000I0320). The Broadbalk Wheat Experiment is part of the Rothamsted Long-term Experiments National Capability supported by BBSRC (BBS/E/C/000J0300) and the Lawes Agricultural Trust. LJG and RK were supported by the Hartree National Centre for Digital Innovation, a collaboration between UKRI's Science and Technology Facilities Council and IBM Research Europe. The authors a grateful to four anonymous reviewers for their time taken to provide helpful and encouraging comments on the original version of the manuscript. Their inputs have resulted in a much-improved text.

References

1. Laborde, D., Mamun, A., Martin, W. *et al*. (2021). Agricultural subsidies and global greenhouse gas emissions. *Nat. Commun.* 12, 2601.
2. Bouwman A.F., Boumans L.J.M., Batjes N.H. (2002). Emissions of N_2O and NO from fertilized fields: summary of available measurement data. *Global Biogeochem. Cycles* 16, 1058-1070.
3. Clark M.A., Domingo N.G.G., Colgan K., *et al*., (2020). Global food system emissions could preclude achieving the 1.5° and 2 °C climate change targets. *Science* 370, 705-708.
4. Guenet, B, Gabrielle, B, Chenu, C, *et al*. (2020). Can N_2O emissions offset the benefits from soil organic carbon storage? *Global Change Biol*. 27: 237– 256.
5. Rappoldt C., Crawford J.W. (1999). The distribution of anoxic volume in a fractal model of soil. *Geoderma* 88, 329-347.
6. Letey, J. (1991). The study of soil structure - science or art. *Austr. J. Soil Res.* 29, 699–707.
7. Neal, A.L., Bacq-Labreuil, A., Zhang, X. *et al.* (2020). Soil as an extended composite phenotype of the microbial metagenome. *Sci. Rep.* 10, 10649.
8. Phillips, J.D. (2009). Soils as extended composite phenotypes. *Geoderma* 149, 142–151.

9. Dawkins, R. *The Extended Phenotype. The Gene as the Unit of Selection* (Oxford University Press, Oxford, 1982).
10. Dawkins, R. (2004). Extended Phenotype—but not too extended. A reply to Laland, Turner and Jablonka. *Biol. Philos*. 19, 377–396.
11. Schjønning, P., Munkholm, L.J., Moldrup, P. *et al*., (2002). Modelling soil pore characteristics from measurements of air exchange: the long-term effects of fertilization and crop rotation. *Euro. J. Soil Sci.* 53, 331–339.
12. Bacq-Labreuil, A., Crawford, J., Mooney, S.J., *et al*., (2018). Effects of cropping systems upon the three-dimensional architecture of soil systems are highly contingent upon texture. *Geoderma* 332, 73–83.
13. Arah, J.R.M., Ball, B.C. (1994). A functional model of soil porosity used to interpret measurements of gas diffusion. *Euro. J. Soil Sci.* 45, 135–144.
14. Ball, B.C. (2013). Soil structure and greenhouse gas emissions: a synthesis of 20 years of experimentation. *Euro. J. Soil Sci.* 64, 357-373.
15. Zhang, X., Gregory, A.S., Whalley, W.R., *et al*., (2021). Relationship between soil carbon sequestration and the ability of soil aggregates to transport dissolved oxygen. *Geoderma*, 403, 115370.
16. Zhang, X., Neal, A.L., Crawford, J.W., *et al*., (2021). The effects of long-term fertilizations on soil hydraulic properties vary with scales. *J. Hydrol.* 593, 125890.
17. Tiedje, J.M. (1988). Ecology of denitrification and dissimilatory nitrate reduction to ammonium, in *Biology of Anaerobic Microorganisms*, edited by A. J. B. Zehnder, pp. 179–244, John Wiley, New York.
18. Ramirez K.S., Lauber C.L., Knight R., *et al*., (2010). Consistent effects of nitrogen fertilization on soil bacterial communities in contrasting systems. *Ecology* 91, 3463-3470.
19. Fierer N, Lauber CL, Ramirez KS, *et al*., (2012). Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. *ISME J*. 6, 1007-1017.
20. Geisseler D., Scow K.M. (2014). Long-term effects of mineral fertilizers on soil microorganisms: a review. *Soil Biol. Biochem*. 75, 54-63.
21. Ouyang Y., Evans S.E., Friesen M.L., *et al*., (2018). Effect of nitrogen fertilization on the abundance of nitrogen cycling genes in agricultural soils: a meta-analysis of field studies. *Soil Biol. Biochem*. 127, 71-78.
22. Knapp, C., Dodds, W. K., Wilson, K. C., *et al*., (2009). Spatial heterogeneity of denitrification genes in a highly homogeneous urban stream. *Environ. Sci. Technol*. 43, 4273–4279.
23. Graham, D. W., Trippett, C., Dodds, W. K., *et al*. (2010). Correlations between in situ denitrification activity and *nir*-gene abundances in pristine and impacted prairie streams. *Environ. Poll*. 158, 3225–3229.
24. Tatariw, C., Chapman, E. L., Sponseller, R. A., *et al*., (2013). Denitrification in a large river: consideration of geomorphic controls on microbial activity and community structure. *Ecology* 94, 2249–2262.
25. Marchant, H., Ahmerkamp, S., Lavik, G. *et al*. (2017). Denitrifying community in coastal sediments performs aerobic and anaerobic respiration simultaneously. *ISME J*. 11, 1799–1812.
26. Petersen, D.G., Blazewicz, S.J., Firestone, M., *et al*., (2012), Abundance of microbial genes associated with nitrogen cycling as indices of biogeochemical process rates across a vegetation gradient in Alaska. *Environ. Microbiol.* 14, 993-1008.
27. Rocca, J., Hall, E., Lennon, J. *et al*. (2015). Relationships between protein-encoding gene abundance and corresponding process are commonly assumed yet rarely observed. *ISME J*. 9, 1693–1699.
28. Chadwick D.R., Cárdenas L., Misselbrook T.H., *et al*., (2014). Optimizing chamber methods for measuring nitrous oxide emissions from plot-based agricultural experiments. *Euro. J. Soil Sci*. 65, 295-307.
29. Raun W.R., Johnson G.V. (1999). Improving nitrogen use efficiency for cereal production. *Agron. J*. 91, 357-363.
30. Erisman JW, Leach A, Bleeker A, *et al*., (2018). An integrated approach to a nitrogen use efficiency (NUE) indicator for the food production–consumption chain. *Sustainability* 10, 925

31. Cárdenas LM, Bhogal A, Chadwick DR, *et al*., (2019). Nitrogen use efficiency and nitrous oxide emissions from five UK fertilised grasslands. *Sci. Total Environ.* 661, 696-710.
32. Jenkinson, D. S. (1977). The nitrogen economy of the Broadbalk experiments. I. Nitrogen balance in the experiments. *Rothamsted Experimental Station Report for 1976*, Part 2, pp. 103-109.
33. Powlson D.S., Pruden G., Johnston A.E., *et al*. (1986). The nitrogen cycle in the Broadbalk Wheat Experiment: recovery and losses of ^{15}N-labelled fertilizer applied in spring and inputs of nitrogen from the atmosphere. *J. Agric. Sci., Camb.* 107, 591-609.
34. Addiscott, T., Powlson, D. (1992). Partitioning losses of nitrogen fertilizer between leaching and denitrification. *J. Agric. Sci., Camb*. 118, 101-107.
35. Skinner C., Gattinger A., Muller A., *et al*., (2014). Greenhouse gas fluxes from agricultural soils under organic and non-organic management — a global meta-analysis. *Sci. Total Environ*. 468-469, 553-563.
36. Körschens M., Albert E., Armbruster M., *et al*., (2013). Effect of mineral and organic fertilization on crop yield, nitrogen uptake, carbon and nitrogen balances, as well as soil organic carbon content and dynamics: results from 20 European long-term field experiments of the twenty-first century. *Arch. Agron. Soil Sci.* 59, 1017-1040.
37. Oelofse M., Markussen B., Knudsen L., *et al*., (2015). Do soil organic carbon levels affect potential yields and nitrogen use efficiency? An analysis of winter wheat and spring barley field trials. *Euro. J. Agron*. 66, 62-73.
38. Hijbeek R., van Ittersum M.K., ten Berge H.F.M., *et al*., (2017). Do organic inputs matter – a meta-analysis of additional yield effects for arable crops in Europe. *Plant Soil* 411, 293-303.
39. Huang J., Hartemink A.E., Kucharik C.J. (2021). Soil-dependent response of US crop yields to climate variability and depth to groundwater. *Agric. Syst*. 190, 103085
40. Kane D.A., Bradford M.A., Fuller E., *et al*., (2021). Soil organic matter protects US maize yields and lowers insurance payouts under drought. *Environ. Res. Lett*. 16, 044018.
41. Macholdt J., Piepho H.-P., Honermeier B., *et al*., (2020). The effects of cropping sequence, fertilization and straw management on the yield stability of winter wheat (1986-2017) in the Broadbalk Wheat Experiment, Rothamsted, UK. *J. Agric. Sci.* 158, 65-79.
42. Powlson, D. S., Poulton, P. R., Glendining, M., *et al*. (2022). Is it possible to attain the same soil organic matter content in arable agriculture soils as under natural vegetation? *Outlook on Agriculture.* 51, 91-104.
43. Guenet, B, Gabrielle, B, Chenu, C, *et al*. (2020). Can N_2O emissions offset the benefits from soil organic carbon storage? *Global Change Biol*. 27, 237– 256.
44. Lugato, E., Leip, A., Jones, A. (2018). Mitigation potential of soil carbon management overestimated by neglecting N_2O emissions. *Nature Clim. Change* 8, 219–223.
45. Macdonald, A., Poulton, P., Clark, I., *et al*. (2018). *Guide to the Classical and Other Long-term experiments, Datasets and Sample Archive. Rothamsted Research*, 57 pp.
46. Vogel, H.-J., Weller, U. & Schlüter, S. (2010). Quantification of soil structure based upon Minkowski functions. *Comput. Geosci.* 36, 1236–1245.
47. Zhang X., Crawford J.W., Flavel R.J., *et al*. (2016). A multi-scale Lattice Boltzmann model for simulating solute transport in 3D X-ray micro-tomography images of aggregated porous materials. *J. Hydrol.* 541, 1020-1029.
48. Li Z., Zhang X., Wang D.I, *et al*. (2018). Direct methods to calculate the mass exchange between solutes inside and outside aggregates in macroscopic model for solute transport in aggregated soil. *Geoderma*, 320, 126-135.
49. Harrison, D.E.F., Pirt, S.J. (1967). The influence of dissolved oxygen concentration on the respiration and glucose metabolism of *Klebsiella aerogenes* during growth. *J. Gen. Microbiol*. 46, 193–211.
50. Bolger, A. M., Lohse, M., Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
51. Buchfink, B., Xie, C., Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Meth*. 12, 59–60.

52. Huson, D.H., Beier S., Flade I., *et al*. (2016). MEGAN Community Edition—interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput. Biol*. 12(6), e1004957.
53. Kanehisa, M., Sato, Y., Kawashima, M., *et al*., (2015). KEGG as a reference resource for gene and protein annotation. *Nucl. Acids Res.* 44, D457–D462.
54. Bateman, E.J., Baggs, E.M. (2005). Contributions of nitrification and denitrification to N_2O emissions from soils at different water-filled pore space. *Biol. Fert. Soils* 41, 379–388.
55. Liang D., Robertson G.P. (2021). Nitrification is a minor source of nitrous oxide (N_2O) in an agricultural landscape and declines with increasing management intensity. *Global Change Biol.* 7, 5599–5613.
56. Cowan N.J., Famulari D., Levy P.E., *et al*. (2014). Investigating uptake of N_2O in agricultural soils using a high-precision dynamic chamber method. *Atmos. Meas. Tech. Discuss*. 7, 8125–8147.

Figure and Table legends

Figure 1. Trends in ***A*** soil organic carbon (SOC) and ***B*** soil total nitrogen (N_{tot}) stocks in unmanaged woodland and grassland soils, and arable soils of the Broadbalk Winter Wheat Experiment receiving farmyard manure (FYM) annually, or inorganic fertiliser at rates of 240 kg-N ha^{-1} yr^{-1} as ammonium nitrate combined with phosphorus and potassium (^{240}NPK), 144 kg-N ha^{-1} yr^{-1} combined with phosphorus, and potassium (^{144}NPK), 192 kg-N ha^{-1} yr^{-1} with potassium but no phosphorus (^{192}NK) and soil which has received no fertilisation (nil). Soil management is summarised in Table I.

Figure 2. Contrasting long-term soil management results in quantitatively different process-form states. ***A*** - Geometric mean functional relationship between soil organic carbon (SOC) and total nitrogen (N_{tot}) spanning the years 1843 – 2015 measured in farmyard manure (FYM)-amended, inorganically-fertilised (^{240}NPK, ^{144}NPK, ^{192}NK) and unfertilised (nil) arable soils of the Broadbalk Winter Wheat Experiment, and unmanaged (woodland and grassland) soils. Slope = 0.088 (bootstrapped 95% confidence interval, 0.085 – 0.091), $t = 60.7$, $p = 6.5\text{x}10^{-63}$; $\rho = 0.990$, $t = 60.1$, $p = 1.3\text{x}10^{-62}$. ***B*** – FYM-amended, ^{144}NPK, ^{192}NK and PK arable soils, and woodland and grassland unmanaged soils are described by a combination of the connectivity of pore space, established from X-ray computed tomography, and simulated saturated hydraulic conductivity (K, in units of μm s^{-1})—a measure of capacity, representing the maximum potential movement of resources through pore networks to organisms. Data point size is proportional to SOC stocks (Mg ha^{-1}) in each soil, shown in Supplementary Fig 1A. ***C*** - Process-form states control anoxia within the same soils. Low-SOC, low-connected porosity soils (inorganically fertilised and unfertilised arable soils ^{144}NPK, ^{192}NK and nil) contain large volumes of anoxic microsites. Across a range of matric potential (ψ_m), the predicted volume of anoxic sites is consistently larger in these soils than high-SOC, high-connected porosity soil (FYM-amended arable, and grassland and woodland soils). Anoxic pore space was modelled under conditions of relatively low microbial respiration, $k' = 8.5\text{x}10^{-3}$. Key shown in *A* relates to all three figures.

Figure 3. Nitrogen metabolism-associated gene assemblages in farmyard manure (FYM)-amended, inorganically fertilised (^{144}NPK, ^{192}NK and PK) arable of the Broadbalk Winter Wheat Experiment, and unmanaged (woodland and grassland) soils determined from shotgun metagenomics. ***Left*** – heatmap representation and bi-hierarchical clustering of gene orthologs (vertical clustering) and soil management (horizontal clustering) based upon centred-log ratio scaled ortholog relative abundance. Genes are identified by their KEGG ortholog (K) number and gene name. The largest $\log_2$-fold difference between soils is shown together with a measure of significance of the difference determined using Wald's test and Benjamini-Hochberg false discovery rate correction (p_{adj}). ***Right*** – association of individual genes with nitrogen metabolism-related functional units of gene sets associated with metabolic pathways (modules). Closed circles indicate genes for which a significant

difference in abundance between soils was identified, open circles indicate genes for which no significant difference was observed. The KEGG ontology does not discriminate between the genes *amoABC* associated with ammonia oxidation and *pmoABC* associated with methane oxidation.

Figure 4. Adjusted mean estimates of nitrous oxide (N_2O) emissions from farmyard manure (FYM)-amended and inorganically fertilised (^{240}NPK, ^{144}NPK, ^{192}NK and PK) soils of the Broadbalk Winter Wheat Experiment measured between April and November 2019. Error bars represent the standard error of the mean, *t* and probabilities (*p*) indicate the results of Holm-Šidák *a priori* contrasts of FYM, ^{144}NPK, ^{192}NK and PK mean N_2O emissions to ^{240}NPK emissions.

Figure 5. Nonequilibrium nitrogen use efficiency of farmyard manure (FYM)-amended and inorganically fertilised (^{240}NPK, ^{192}NPK, ^{144}NPK and PK) soils of the Broadbalk Winter Wheat Experiment between 2000 and 2015 indicates divergent allocation of ***A*** – absolute nitrogen inputs to arable systems and ***B*** - nitrogen inputs supporting 1 Mg of grain production, between grain, straw, soil and lost pools. Small increases in nitrogen soil stocks observed for inorganic plots could be due to measurement error, as these systems are assumed to have reached steady states. ***C*** – Grain yields for inorganically fertilised and FYM-amended soils for the same period. Box plots show interpolated 25% and 75% quartiles, the median is represented as a horizontal line. Whiskers indicate minimum and maximum values. Each treatment mean is represented by a X, the line connects treatment means. Results of analysis of variance (ANOVA) are shown together with estimates of treatment effect size (ω^2) upon yield, calculated as $(SS_{bg} - df_{bg} \cdot MS_{wg})/(SS_{total}+MS_{wg})$ where SS indicates the sum of squares, df the degree of freedom, MS the mean square and the subscripts bg and wg indicate between-group and within-group respectively. *Post hoc* pairwise comparisons indicate that the mean PK grain yield was significantly lower than the means of all other treatments (smallest difference, $Q = 11.7$, $p = 2\text{x}10^{-12}$). FYM mean grain yield was significantly lower than the mean yields of both the ^{192}NPK and ^{240}NPK treatments (smallest difference, $Q = 4.2$, $p = 0.033$). There was no significant difference between the mean grain yields of FYM and ^{144}NPK treatments.

Table I. Fertility management associated with arable soils of the Broadbalk Winter Wheat Experiment and associated unmanaged soils used in various aspects of this study.

Table II. Topology-related parameters derived directly from binary images generated from X-ray computed tomography of aggregates from farmyard manure (FYM)-amended and inorganically fertilized (^{144}NPK, ^{192}NK and PK) arable soils of the Broadbalk Winter Wheat Experiment and unmanaged woodland and grassland (total and connected porosity) and estimates of permeability (*k*), effective diffusion coefficient of oxygen (D_e') which is normalised by dividing the diffusion coefficient of oxygen in water without pore constraints, and hydraulic conductivity (*K*) of the soil pore networks derived from lattice Boltzman simulation (and where $K = gk/\nu$ where g is the gravitational constant, k the permeability, and ν the viscosity of water). The mean associated with each parameter is provided together with the standard error of the mean in parentheses. Results of analysis of variance (ANOVA) are shown together with estimates of treatment effect size (ω^2) for each parameter, calculated as $(SS_{bg} - df_{bg} \cdot MS_{wg})/(SS_{total}+MS_{wg})$ where SS indicates the sum of squares, df the degree of freedom, MS the mean square and the subscripts bg and wg indicate between-group and within-group respectively.

Supplementary Figure 1. ***A*** - Soil organic carbon (SOC) stocks for the years 1966 – 2015 measured in farmyard manure (FYM)-amended, inorganically fertilised (^{240}NPK, ^{144}NPK, ^{192}NK and PK) and unfertilised (nil) arable soils of the Broadbalk Winter Wheat Experiment and unmanaged grassland soil. ***B*** - Soil total nitrogen (N_{tot}) stocks for the same period and treatments. Box plots indicate the interpolated 25% and 75% quartiles with the median represented as a horizontal line. Whiskers indicate minimum and maximum values. Results of analysis of variance (ANOVA) are shown together with estimates of treatment effect size (ω^2) upon SOC and N_{tot}, calculated as $(SS_{bg} - df_{bg} \cdot MS_{wg})/(SS_{total}+MS_{wg})$ where SS indicates the sum of squares, df the degree of freedom, MS the mean square and the subscripts bg and wg indicate between-group and within-group respectively.

Supplementary Figure 2. Mean estimates (n = 3) of nitrous oxide (N_2O) emissions from farmyard manure (FYM)-amended and inorganically fertilised (^{240}NPK, ^{144}NPK, ^{192}NK and PK) soils of the Broadbalk Winter Wheat Experiment measured in 2019. Emissions were measured discontinuously on the dates shown using in-field static chambers. Daily precipitation measured on a 254-mm diameter rain gauge located 500 m east of the field experiment together with the potential soil moisture deficit (PSMD) are also shown.

Supplementary Figure 3. Nonequilibrium nitrogen use efficiency of farmyard manure (FYM)-amended and inorganically fertilised (^{240}NPK, ^{192}NPK, ^{144}NPK and PK) treatments of the Broadbalk Winter Wheat Experiment between 2000 and 2015. The proportion of total nitrogen inputs derived from ammonium nitrate fertiliser or FYM, associated with seed grain and atmospheric deposition allocated to harvested grain and straw, and soil and lost pools.

Figure 1.

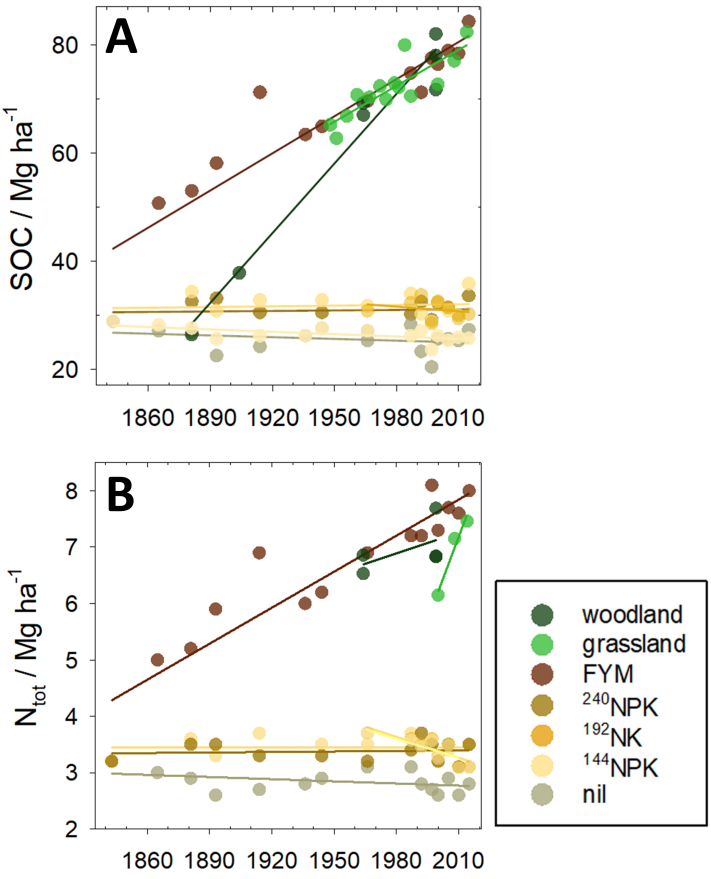
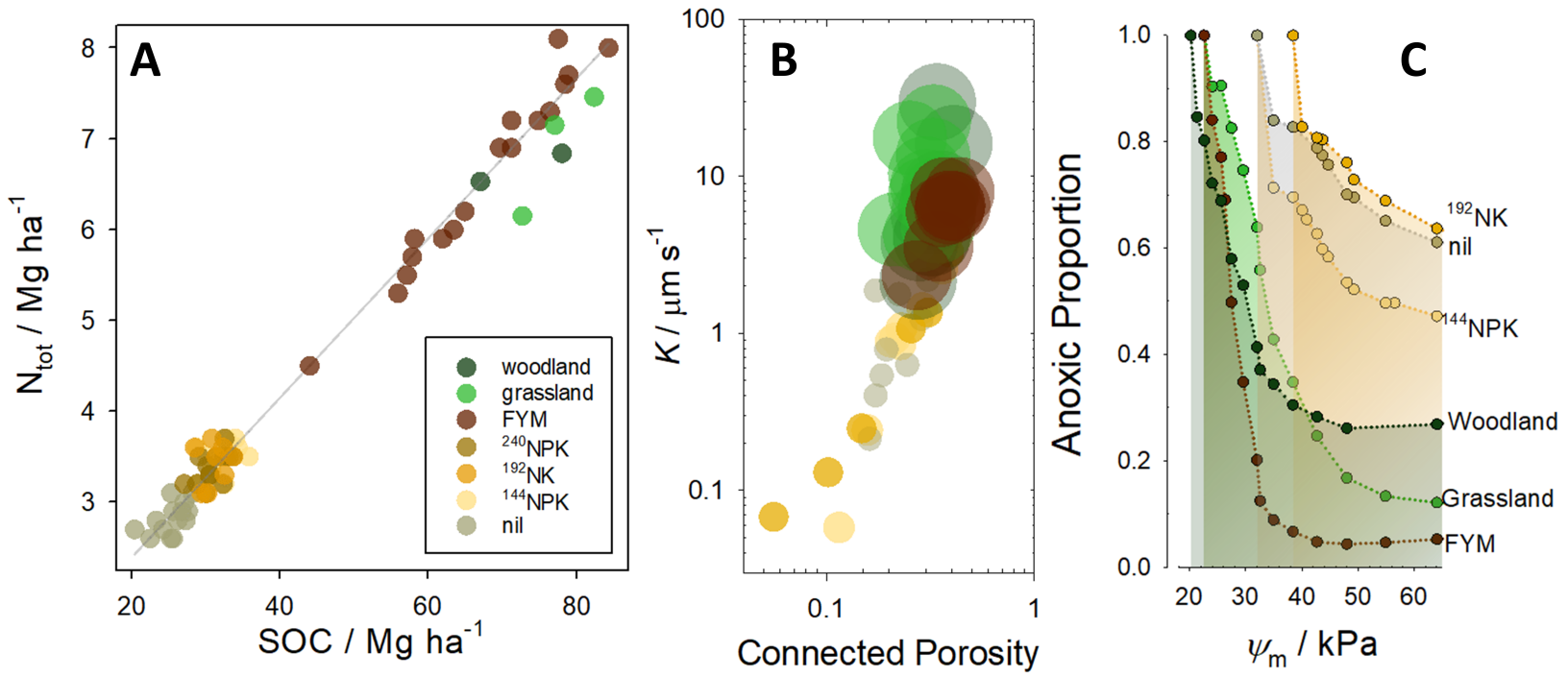

A
B
SOC / Mg ha-1
Ntot / Mg ha-1
80
60
40
20
8
7
6
5
4
3
2
1860 1890 1920 1950 1980 2010
woodland
grassland
FYM
240NPK
192NK
144NPK
nil

Figure 2.


A
B
C
N_{tot} / Mg ha^{-1}
SOC / Mg ha^{-1}
woodland
grassland
FYM
^{240}NPK
^{192}NK
^{144}NPK
nil
K / µm s^{-1}
Connected Porosity
Anoxic Proportion
ψ_m / kPa
^{192}NK
nil
^{144}NPK
Woodland
Grassland
FYM

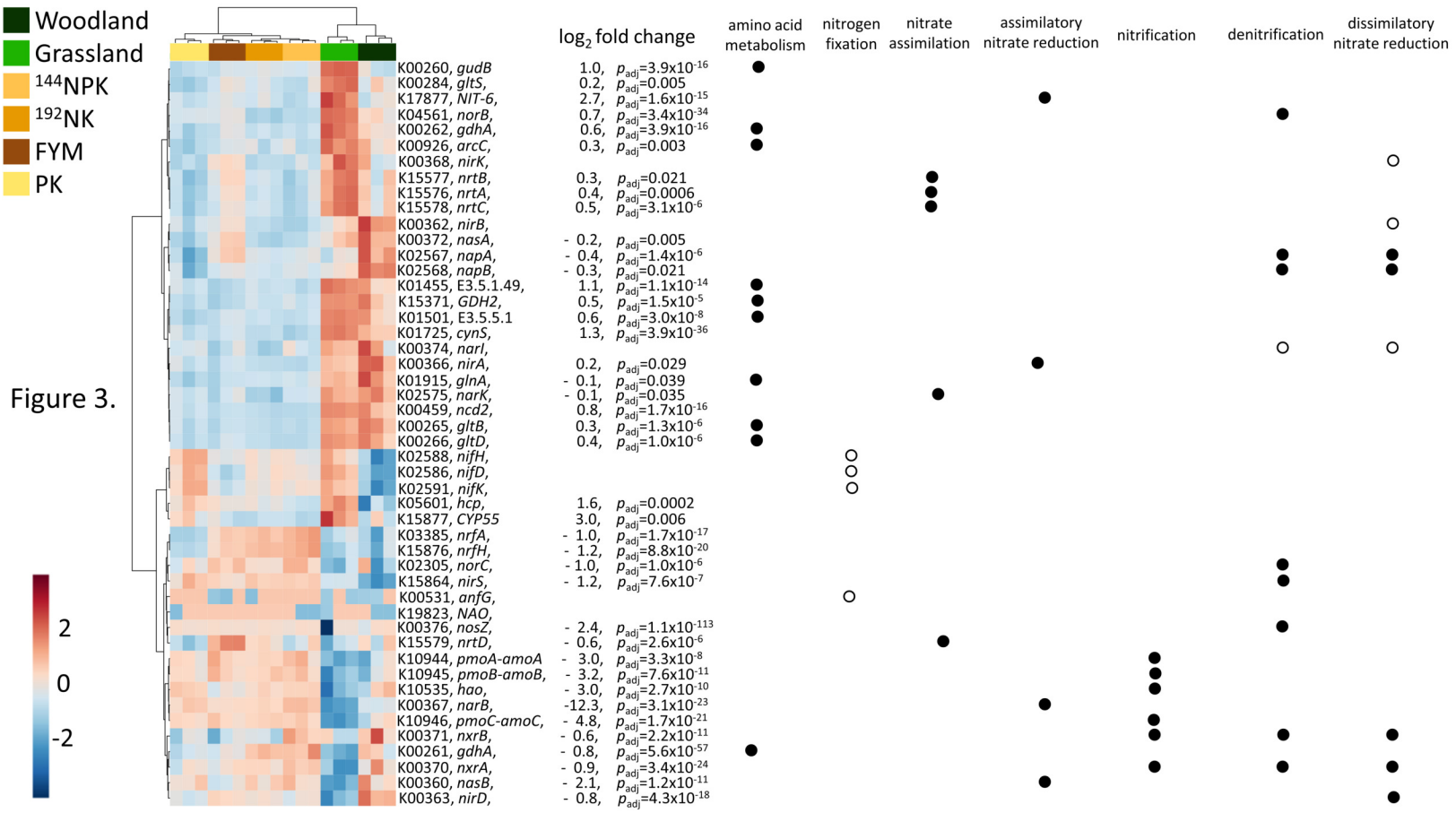

Woodland
Grassland
144NPK
192NK
FYM
PK
2
0
-2
log2 fold change
amino acid metabolism
nitrogen fixation
nitrate assimilation
assimilatory nitrate reduction
nitrification
denitrification
dissimilatory nitrate reduction
K00260, gudB 1.0, padj=3.9x10-16
K00284, gltS, 0.2, padj=0.005
K17877, NIT-6, 2.7, padj=1.6x10-15
K04561, norB, 0.7, padj=3.4x10-34
K00262, gdhA, 0.6, padj=3.9x10-16
K00926, arcC, 0.3, padj=0.003
K00368, nirK,
K15577, nrtB, 0.3, padj=0.021
K15576, nrtA, 0.4, padj=0.0006
K15578, nrtC, 0.5, padj=3.1x10-6
K00362, nirB,
K00372, nasA, - 0.2, padj=0.005
K02567, napA, - 0.4, padj=1.4x10-6
K02568, napB, - 0.3, padj=0.021
K01455, E3.5.1.49, 1.1, padj=1.1x10-14
K15371, GDH2, 0.5, padj=1.5x10-5
K01501, E3.5.5.1 0.6, padj=3.0x10-8
K01725, cynS, 1.3, padj=3.9x10-36
K00374, narI,
K00366, nirA, 0.2, padj=0.029
K01915, glnA, - 0.1, padj=0.039
K02575, narK, - 0.1, padj=0.035
K00459, ncd2, 0.8, padj=1.7x10-16
K00265, gltB, 0.3, padj=1.3x10-6
K00266, gltD, 0.4, padj=1.0x10-6
K02588, nifH,
K02586, nifD,
K02591, nifK,
K05601, hcp, 1.6, padj=0.0002
K15877, CYP55 3.0, padj=0.006
K03385, nrfA, - 1.0, padj=1.7x10-17
K15876, nrfH, - 1.2, padj=8.8x10-20
K02305, norC, - 1.0, padj=1.0x10-6
K15864, nirS, - 1.2, padj=7.6x10-7
K00531, anfG,
K19823, NAO,
K00376, nosZ, - 2.4, padj=1.1x10-113
K15579, nrtD, - 0.6, padj=2.6x10-6
K10944, pmoA-amoA - 3.0, padj=3.3x10-8
K10945, pmoB-amoB, - 3.2, padj=7.6x10-11
K10535, hao, - 3.0, padj=2.7x10-10
K00367, narB, -12.3, padj=3.1x10-23
K10946, pmoC-amoC, - 4.8, padj=1.7x10-21
K00371, nxrB, - 0.6, padj=2.2x10-11
K00261, gdhA, - 0.8, padj=5.6x10-57
K00370, nxrA, - 0.9, padj=3.4x10-24
K00360, nasB, - 2.1, padj=1.2x10-11
K00363, nirD, - 0.8, padj=4.3x10-18

Figure 3.

Figure 4.

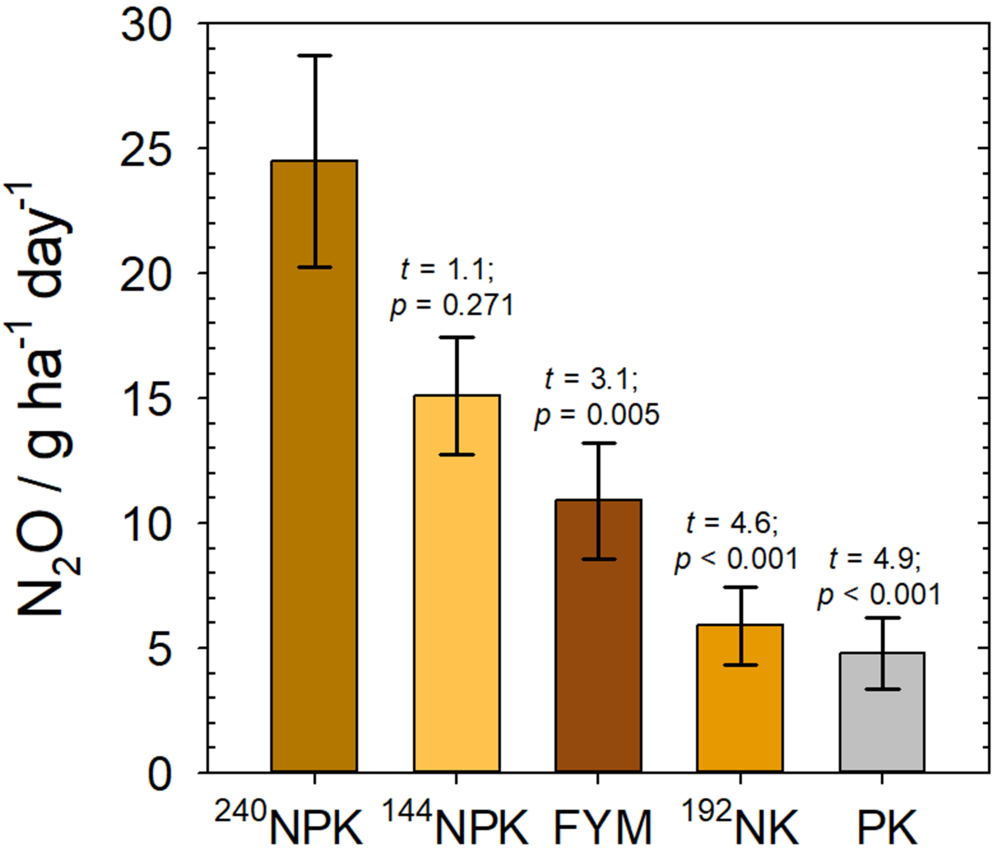

30
25
20
15
10
5
0
N_2O / g ha^{-1} day^{-1}
t = 1.1;
p = 0.271
t = 3.1;
p = 0.005
t = 4.6;
p < 0.001
t = 4.9;
p < 0.001
240NPK
144NPK
FYM
192NK
PK

Figure 5.

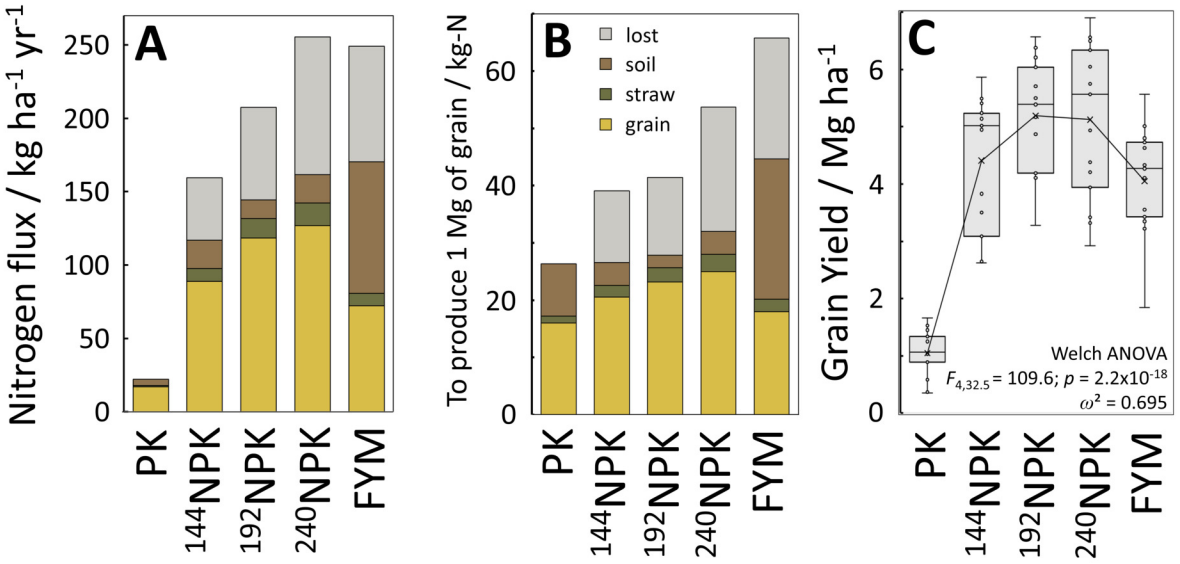

A
Nitrogen flux / kg ha-1 yr-1
250
200
150
100
50
0
PK
144NPK
192NPK
240NPK
FYM
B
lost
soil
straw
grain
To produce 1 Mg of grain / kg-N
60
40
20
0
C
Grain Yield / Mg ha-1
6
4
2
0
Welch ANOVA
F4,32.5 = 109.6; p = 2.2x10-18
ω² = 0.695

Table I.

Treatment	Annual fertility management	Designation
Arable, composted farmyard manure	35 Mg ha^{-1} since 1843	FYM
Arable, 240 kg-N ha^{-1} inorganic fertiliser[a]	240 kg-N ha^{-1} since 1985; (144 kg-N ha^{-1} between 1852 and 1984)	^{240}NPK
Arable, 192 kg-N ha^{-1}, inorganic fertiliser[a]	192 kg-N ha^{-1} since 1968	^{192}NPK
Arable, 144 kg-N ha^{-1} inorganic fertiliser[a]	144 kg-N ha^{-1} since 1852	^{144}NPK
Arable, inorganic fertiliser[a], no phosphorus	No triple superphosphate additions since 1968; (96 kg-N ha^{-1} between 1968 and 2001; 192 kg-N ha^{-1} since)	^{192}NK
Arable, inorganic fertiliser[a], no nitrogen	No ammonium nitrate additions since 1852	PK
Arable, unfertilised	No fertiliser additions since 1852	nil
Grassland	None, grassland since 1838	grassland
Woodland	None, woodland since 1882	woodland

[a] – unless stated otherwise, inorganic fertiliser included inorganic nitrogen as ammonium nitrate, 90 kg-K ha^{-1} as K_2SO_4, 35 kg-P ha^{-1} as triple superphosphate [$Ca(H_2PO_4)_2 \cdot H_2O$], and 12 kg-Mg ha^{-1} as kieserite ($MgSO_4 \cdot H_2O$)

Table II.

	Total Porosity (P_t) / %	**Connected Porosity (P_c) / %**	**Permeability (k) / µm²**	**Diffusion Coefficient (D_e')**	**Hydraulic Conductivity (K) / µm s^{-1}**
	Welch ANOVA; $F_{5,20.8} = 14.1$, $p = 4.5\text{x}10^{-6}$ $\omega^2 = 0.467$	Welch ANOVA; $F_{5,20.7} = 14.2$, $p = 4.3\text{x}10^{-6}$ $\omega^2 = 0.459$	Welch ANOVA; $F_{5,21} = 36.5$, $p = 1.4\text{x}10^{-9}$ $\omega^2 = 0.775$	ANOVA; $F_{5,47} = 19.5$, $p = 1.8\text{x}10^{-10}$ $\omega^2 = 0.636$	Welch ANOVA; $F_{5,21} = 15.8$, $p = 2.1\text{x}10^{-6}$ $\omega^2 = 0.361$
woodland	33.1 (1.3)	32.8 (1.4)	0.396 (0.131)	0.278 (0.020)	8.86 (2.94)
grassland	31.0 (1.2)	30.8 (1.2)	1.13 (0.089)	0.254 (0.015)	9.61 (1.61)
FYM	37.7 (1.6)	37.5 (1.6)	0.265 (0.028)	0.335 (0.023)	5.92 (0.62)
^{144}NPK	23.6 (2.4)	22.4 (2.7)	0.045 (0.012)	0.135 (0.016)	1.01 (0.27)
^{192}NK	24.9 (4.5)	23.4 (5.1)	0.065 (0.027)	0.151 (0.032)	1.45 (0.60)
PK	22.6 (1.7)	21.6 (1.8)	0.048 (0.011)	0.129 (0.014)	1.07 (0.24)