



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/125934/>

Version: Accepted Version

Article:

Blaud, A., van der Zaan, B., Menon, M. et al. (2018) The abundance of nitrogen cycle genes and potential greenhouse gas fluxes depends on land use type and little on soil aggregate size. *Applied Soil Ecology*, 125. pp. 1-11. ISSN: 0929-1393

<https://doi.org/10.1016/j.apsoil.2017.11.026>

Article available under the terms of the CC-BY-NC-ND licence
(<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Highlights:

- 1) Land use is the main factor explaining N cycle genes abundance and GHG fluxes
- 2) Soil aggregates size is a minor factor explaining N genes abundance and GHG fluxes
- 3) Cropland showed the lowest abundance for bacteria, fungi, *nifH*, *narG*, *nirS* and *nosZ*
- 4) Effect of aggregate sizes on N genes abundance was only found in forest sites
- 5) Aggregates 0.5 – 1.0 mm showed the highest N functional genes abundance in forest sites

1 **The abundance of nitrogen cycle genes and potential greenhouse gas fluxes depends on**
2 **land use type and little on soil aggregate size**

3
4 Aimeric Blaud^{a, 1*}, Bas van der Zaan^b, Manoj Menon^{a, 2}, Georg J. Lair^{c, d}, Dayi Zhang^{a, 3}, Petra
5 Huber^c, Jasmin Schiefer^c, Winfried E.H. Blum^c, Barbara Kitzler^e, Wei E.Huang^{a, 4}, Pauline van
6 Gaans^b, and Steve Banwart^a

7
8 ^a Department of Civil and Structural Engineering, Kroto Research Institute, University of
9 Sheffield, Broad Lane, Sheffield S3 7HQ, United Kingdom.

10 ^b Deltares, Subsurface and Groundwater Systems, Princetonlaan 6-8, 3508 AI Utrecht, the
11 Netherlands.

12 ^c University of Natural Resources and Life Sciences (BOKU), Institute of Soil Research, Vienna,
13 Peter-Jordan-Str. 82, 1190 Vienna, Austria.

14 ^d University of Innsbruck, Institute of Ecology, Sternwartestr. 15, 6020 Innsbruck, Austria.

15 ^e Department of Forest Ecology and Soil, Soil Ecology, Federal Research Centre for Forests,
16 Seckendorff-Gudent-Weg 8, 1131 Vienna Austria.

17
18 *Corresponding Author.

19 E-mail address: aimeric.blaud@gmail.com

20
21 ¹ Current address: Sustainable Agriculture Science Department, Rothamsted Research,
22 Harpenden, Hertfordshire AL5 2JQ, UK.

23 ² Current address: Department of Geography, University of Sheffield, Winter street, Sheffield S10
24 2TN, United Kingdom.

25 ³ Current address: School of Environment, Tsinghua University, Beijing 200084, PR China.

26 ⁴ Current address: Department of Engineering Science, University of Oxford, Parks Road, Oxford
27 OX1 3PJ, UK.

28

29 **Abstract**

30 Soil structure is known to influence microbial communities in soil and soil aggregates
31 are the fundamental ecological unit of organisation that support soil functions. However, still
32 little is known about the distribution of microbial communities and functions between soil
33 aggregate size fractions in relation to land use. Thus, the objective of this study was to
34 determine the gene abundance of microbial communities related to the nitrogen cycle and
35 potential greenhouse gas (GHG) fluxes in six soil aggregate sizes (0-0.25, 0.25-0.5, 0.5-1.0, 1-2,
36 2-5, 5-10 mm) in four land uses (i.e. grassland, cropland, forest, young forest). Quantitative-PCR
37 (Q-PCR) was used to investigate the abundance of bacteria, archaea and fungi, and functional
38 guilds involved in N-fixation (*nifH* gene), nitrification (bacterial and archaeal *amoA* genes) and
39 denitrification (*narG*, *nirS*, and *nosZ* genes). Land use leads to significantly different
40 abundances for all genes analysed, with the cropland site showing the lowest abundance for all
41 genes except *amoA* bacteria and archaea. In contrast, not a single land use consistently showed
42 the highest gene abundance for all the genes investigated. Variation in gene abundance between
43 aggregate size classes was also found, but the patterns were gene specific and without common
44 trends across land uses. However, aggregates within the size class of 0.5 – 1.0 mm showed high
45 bacterial 16S, *nifH*, *amoA* bacteria, *narG*, *nirS* and *nosZ* gene abundance for the two forest sites
46 but not for fungal ITS and archaeal 16S. The potential GHG fluxes were affected by land use but
47 the effects were far less pronounced than for microbial gene abundance, inconsistent across
48 land use and soil aggregates. However, few differences in GHG fluxes were found between soil
49 aggregate sizes. From this study, land use emerges as the dominant factor that explains the
50 distribution of N functional communities and potential GHG fluxes in soils, with less pronounced
51 and less generalized effects of aggregate size.

52

53 **Keywords:** Quantitative-PCR; nitrogen-fixation; nitrification; denitrification; soil aggregates;
54 land use

55

56 **1. Introduction**

57

58 Soil is a complex and heterogeneous matrix made up of an intricate organisation of
59 pores filled with water and gas, mineral particles, and organic matter influencing the
60 microorganisms that live within. Soil aggregates are essential for soil fertility (Amézqueta, 1999;
61 Bronick and Lal, 2005) and some fertile soils have been described as soils dominated by 0.25 –
62 10 mm soil crumbs (Shein, 2005). The vast variation in the size of aggregates, as well as their
63 physico-chemical properties provides a huge diversity of microhabitats for microorganisms
64 influencing carbon and nutrients dynamics within the soil. This study starts from the premise
65 that soil aggregates are a fundamental ecological unit of organisation that support soil functions.
66 These soil functions include biomass production, soil water retention and transmission, nutrient
67 transformation, contaminant attenuation, C and N, P, K sequestration, and a major terrestrial
68 pool of genetic diversity. The microbial community has been found to vary with the size of soil
69 aggregates, and to be linked to the specific environmental conditions in the different sizes of
70 aggregates. Previous studies showed differences in microbial community structure, diversity
71 and abundance/biomass between soil aggregates of different size, which was correlated to the
72 quality of organic matter available (Blaud et al., 2012; Davinic et al., 2012), the size of the pores
73 (Kravchenko et al., 2014) or tillage (Helgason et al., 2010).

74 Although the distribution of microbial communities in soil aggregates has been studied,
75 much less is known about the distribution of the microbial functional guilds among soil
76 aggregates and how their sizes influence microbial functions. The size of soil aggregates in
77 relation to their porosity (i.e. size and number of pores) was found to affect the GHG fluxes, with
78 CO₂ emissions found to be higher in microaggregates (< 0.25 mm) than in macroaggregates (>
79 0.25 mm) in cropland sandy loam soil (Sey et al., 2008; Mangalassery et al., 2013). Similar
80 results were found for CH₄ in cropland sandy loam and clay loam soil (Mangalassery et al.,
81 2013), but the contrary was found in paddy rice soil (Ramakrishnan et al., 2000). Only a few

82 studies have investigated specific microbial functional guilds such as N fixation (Mendes and
83 Bottomley, 1998; Poly et al., 2001; Chotte et al., 2002; Izquierdo and Nüsslein, 2006) and
84 denitrifiers (Beauchamp and Seech, 1990; Lensi et al., 1995) in soil aggregates. The biomass and
85 composition of diazotrophs varies with the size of soil aggregates which was correlated with
86 total C and N, and soil texture (Poly et al., 2001; Izquierdo and Nüsslein, 2006). Aggregates
87 within size classes 0.6 – 2.0 mm and < 0.075 mm (from tundra, pasture and forest) were found
88 to have the highest diazotroph richness (Izquierdo and Nüsslein, 2006) and microaggregates (<
89 0.25 mm) to host between 30% and 90% of the diazotrophic population (Mendes and
90 Bottomley, 1998; Chotte et al., 2002). In contrast, denitrifiers were found to occur mainly in
91 microaggregates, where nearly 90% of the potential denitrification activity can occur (Lensi et
92 al., 1995). Hence, the diazotroph and denitrifier communities seem to exploit specific and
93 different anaerobic niches within different soil aggregate size classes, although the drivers of
94 these communities in different soil aggregate sizes remains unclear.

95 The type of land use and management directly influences the physico-chemical
96 properties of soil aggregates as well as the distribution of microbial communities, their
97 functions and resulting nutrient transformations and GHG fluxes. For example, the soil
98 aggregates turnover rate is increased by soil tillage (Six et al., 2004), which decreases the C
99 storage within the aggregates (Bossuyt et al., 2002), but can also decrease N₂O fluxes (Ball,
100 2013). Furthermore, the type of vegetation and input of organic manure influence the aggregate
101 size distribution and the contents of organic C and N within soil aggregates (Pinheiro et al., 2004;
102 Six et al., 2004; An et al., 2010). Subsequently, bacterial and fungal community composition was
103 found to differ between land use types (Lauber et al., 2008) and also microbial activity such as
104 nitrification (Hayden et al., 2010).

105 The above leads to the overarching hypothesis that in conjunction with land use,
106 different microbial functions are preferentially hosted or fostered by specific size classes of
107 aggregates. The specific objectives of the current study were: i) to assess the difference in
108 microbial genes abundance between different soil aggregate size classes and bulk soil from

109 different land uses, ii) to assess the difference in greenhouse gases fluxes between soil
110 aggregate sizes classes and bulk soil from different land uses, iii) to identify possible
111 relationships between microbial gene abundances, potential GHG fluxes and the physico-
112 chemical characteristics of the soil aggregates.

113

114 **2. Material and methods**

115

116 **2.1 Study area**

117 The study area is originated from the Critical Zone Observatory Marchfeld/Fuchsenbigl
118 area (Banwart, 2011) located east of Vienna, Austria, in the National Park “Donau-Auen” on a
119 floodplain of the Danube River (Fig. S1). The mean annual temperature in the area is ~9 °C and
120 mean annual precipitation ~550 mm. The study sites are located along a chronosequence
121 starting from a young river island (created <70 years; average inundation frequency: 10 day yr⁻¹)
122 named “young forest”, and sites disconnected from the river through a flood control dike: forest,
123 grassland and cropland. The young forest is impacted by flood events, and covered by “soft-
124 wood” dominated by *Salicetum albae*, while the forest site is covered by “hard-wood”
125 dominated by *Fraxino-Ulmetum* (Schubert et al., 2001), respectively. The grassland site was
126 converted from forest to grassland (presently *Onobrychido viciifoliae-Brometum*) between
127 1809 and 1859 and is currently cut twice a year. The cropland site was grassland before 1781
128 and was converted to intensive cropland in the first half of the 20th century. Cropland site was
129 conventionally managed, with annual tillage and NPK mineral fertilisers. The field is under crop
130 rotation (maize, sugar beet, barley and wheat), with summer wheat the year of the sampling
131 which was shortly harvested before the soil sampling. According to Lair et al. (2009), the topsoil
132 (0-10 cm) of the young forest was deposited after 1986, whereas a topsoil age of approx. 250-
133 350 years on the forest, grassland, and cropland site can be estimated. The soils are classified as
134 Epigleyic Fluvisol (young forest) and Mollic Fluvisols (forest, grassland and cropland; (IUSS
135 Working Group WRB, 2014). The Epigleyic Fluvisol is at least one time of the year impacted by

136 groundwater and is located close to the Danube River. In contrast, the Mollic Fluvisols have no
137 impact of groundwater and are characterized by a fast OC accumulation in the topsoil. In our
138 study area Mollic Fluvisols develop towards a Chernozem.

139

140 ***2.2 Soil sampling and fractionation***

141 The soil sampling was identical at all sites and was performed in September 2011 under
142 dry soil moisture conditions (capillary potential pF 3.8 - 4.0). At each site, three sampling spots
143 (70 x 70 cm) were randomly selected within a circle of about 30 m radius. The soil layer from 5 -
144 10 cm soil depth was sampled to avoid the main rooting zone in grassland and the litter layer in
145 forest sites, focusing on the similar mineral soil layer across sites. The soil samples were
146 manually dry sieved to obtain 6 soil aggregate size classes: < 0.25, 0.25 - 0.5, 0.5 - 1, 1 - 2, 2 - 5,
147 and 5 - 10 mm. The soil fraction > 10 mm was not included in the study as it was composed of a
148 wide range of aggregates and large clumps (100 – 500 g per clump). During dry sieving, visible
149 roots were removed. Sieving continued with freshly excavated soil until ~200 g of soil
150 aggregates was obtained for each aggregate size class. Additional bulk soil samples were
151 collected at each site and sampling spot. Soil aggregate size fractions and bulk soil samples were
152 stored at 4 °C and samples for DNA extraction at -20°C before subsequent analysis. Dry-sieving
153 was chosen over wet-sieving to avoid any bias due to dry/wet cycles with wet-sieving that could
154 have direct effect on GHG emissions (Kaiser et al., 2015). Despite knowing that the sieving
155 method affects the gene abundance quantification, dry-sieving can nonetheless reveal
156 differences in gene abundance between soil aggregate sizes (Blaud et al., 2017).

157

158 ***2.3. DNA extraction and quantitative-PCR***

159 Total nucleic acids were extracted from 0.20 to 0.55 g of fresh soil aggregates from all
160 size classes and from bulk soil samples with PowerSoil® DNA Isolation Kit (Mo-Bio laboratories,
161 Carlsbad, CA, USA) according to manufacturer's instruction, except for the final step where the
162 nucleic acids were eluted in 100 µl of sterile nuclease free water instead of solution C6.

163 Microbial abundance was investigated by Quantitative-PCR (Q-PCR) targeting specific genes or
164 genetic regions. Bacterial and archaeal communities were targeted via the 16S rRNA genes,
165 while the fungal community abundance was investigated by targeting the ITS region. The
166 different communities involved in most steps of the N-cycle were investigated: the nitrogen
167 fixing microorganisms were quantified based on the *nifH* gene; nitrification was investigated by
168 targeting the ammonia oxidising bacteria (AOB) and archaea (AOA) via the *amoA* gene, and
169 denitrifiers were targeted via the *narG* gene coding for the nitrate reductase, the *nirS* gene
170 coding for the nitrite reductase and the *nosZ* gene coding for the nitrous oxide reductase (Table
171 S1).

172 Q-PCR standards for each molecular target were obtained using a 10-fold serial dilution
173 of plasmids carrying a single cloned target gene or relevant part thereof. Standard curve
174 template DNA and the “no template control” (NTC) were amplified in duplicate in the same plate
175 as the environmental samples. Q-PCR amplifications were performed in 25 µl volumes
176 containing 12.5 µl of iQ™ SYBR® Green Supermix (Bio-Rad, Hemel Hempstead, UK), 8.5 µl of
177 nuclease-free water (Ambion, Warrington, UK), 1.25 µl of each primer (10 µM) and 1 µl of
178 template DNA using a CFX96™ Real-Time System (Bio-Rad, Hemel Hempstead, UK).

179 Amplification conditions for all Q-PCR assays are given in the supplementary material and Table
180 S1. The efficiency of the Q-PCR assays was above 90%, except for fungi and AOA (~70%). The r^2
181 were > 0.99, except for *nifH* and *nosZ* genes (~0.97).

182

183 ***2.4. Microbial respiration***

184 Greenhouse gas fluxes from the aggregate size fractions and the bulk soil were
185 measured from field moist bulk soil and soil aggregates (pF 3.8 -4.0; hereafter named “field
186 moisture”) and from moistened samples (40 – 60 % of field capacity) by adding distilled water
187 48 hours before flux measurements started (hereafter named “elevated moisture”). Soil
188 temperature was set to 20 °C. The soil moisture was increased because at the time of soil
189 sampling the soil moisture content was low (pF 3.8-4.0), potentially reducing microbial activity

190 and subsequent GHG fluxes. For full details on the GHG measurements, refer to the
191 supplementary material.

192 Fluxes of CO₂ and NO were measured with a fully automated laboratory measuring
193 system as described in detail by Schindlbacher et al. (2004) and Schaufler et al., (2010). Carbon
194 dioxide was measured with a PP Systems WMA-2 (Amesbury, MA, USA), infrared CO₂ analyser,
195 and NO was measured with a HORIBA APNA-360 (Kyoto, Japan) chemoluminescence NO_x
196 analyser. Determination of N₂O and CH₄ fluxes was done manually by closed chamber technique.
197 The analysis was done immediately after gas sampling by gas chromatography (AGILENT
198 6890N) connected to an automated system sample-injection (AGILENT TECH G1888, Network
199 HEADSPACE-SAMPLER) at an oven temperature of 40 °C. Nitrous oxide was measured by a ⁶³Ni-
200 electron-capture detector and CH₄ by a flame ionization detector.

201

202 ***2.5. Physico-chemical analysis of bulk soil and aggregates***

203 The soil moisture content, organic C, total N, N-NO₃⁻, N-NH₄⁺, P-PO₃⁻⁴, and carbonate
204 concentration, C/N, and soil texture (i.e. sand, silt and clay contents) were measured for each
205 aggregate size class and bulk soil. Three different fractions of soil organic matter (SOM) were
206 determined by simultaneous thermal analysis (STA) according to Barros et al. (2007): labile
207 SOM, stable SOM and refractory SOM. Particle size distribution in the various aggregate size
208 classes as well as the SOM fractions (STA) were measured on one composite sample for each
209 site (i.e. mixture of the 3 replicates at each site). For full details of the methods used, refer to the
210 supplementary material.

211

212 **2.6 Statistical analysis**

213 To test the effects of land use and soil aggregate size on microbial gene abundance, GHG
214 fluxes and soil aggregate characteristics, analyses of variance (ANOVA) were performed with
215 land use and soil aggregate size as factors (3 and 6 degrees of freedom (df) respectively). The
216 normality of the model residuals and the homoscedasticity of the variances were checked before

217 statistical analysis. When one or both of these conditions were not met, the data were log
218 transformed to comply with the conditions. However, if log transformation did not lead to
219 normality or homoscedasticity or could not be applied (presence of negative values for GHG),
220 one-way ANOVA was performed to test the effect of land use within each aggregate size class
221 separately.

222 Similarly, to test the effect of soil moisture level on GHG fluxes for each land use, two-way
223 ANOVA was applied with soil aggregate size and soil moisture level as main factors.

224 To test the effect of aggregate size within each land use on microbial gene abundance,
225 GHG fluxes and soil aggregate characteristics, one-way ANOVA was performed with aggregates
226 size as a factor ($df = 6$) for each land use separately, insuring conditions were met as described
227 previously. When significant ($P < 0.05$) effects were found for ANOVA, the Tukey HSD (honest
228 significant difference) test was used to reveal the significance of the differences between class
229 pairs.

230 In order to get insight into the potential drivers of microbial gene abundances and GHG
231 fluxes, Spearman's rank correlation coefficients ρ ($-1 \leq \rho \leq 1$) were calculated between
232 microbial gene abundance, GHG and soil characteristics, across all the land uses to reveal the
233 factors explaining the differences due to land use, or for each land use to reveal the factors
234 explaining the differences due to soil aggregate size classes. To display the correlations,
235 heatmaps were constructed using the library "gplots" from R software, where colours represent
236 the direction and strength of the correlation.

237 All statistical analyses were performed using R v3.2.1 (R Development Core Team, 2015)
238 and a significance level of $P < 0.05$ was used throughout.

239

240 ***3. Results***

241 ***3.1 Variation in soil aggregates characteristics***

242 The physico-chemical parameters of soil aggregates significantly differed between land
243 use, and between aggregates size classes. The soil aggregate mass distribution showed the same

244 pattern for all the land uses, with the size class 2.0 – 5.0 mm being the most abundant (20 – 40
245 w/w %), and size classes < 0.25 mm the least (< 10%; Fig. S2). Young forest and forest showed
246 significantly higher soil water content for most soil aggregate sizes in comparison to cropland
247 and grassland (Fig. S2). The cropland soil had the lowest soil organic C (SOC) and total N
248 concentrations (~25 and ~1.5 g kg⁻¹ soil, respectively), whereas the grassland soil showed the
249 highest concentrations (~50 and ~3 g kg⁻¹ soil, respectively; Fig. S3). Grassland showed
250 significantly lower N-NO₃⁻ concentration for soil aggregates > 0.5 mm (~10 times) than the
251 other sites, but significantly higher N-NH₄⁺ for the bulk soil (~5 times) and some soil aggregates
252 (Fig. S4). The P-PO₃³⁻ in cropland was significantly lower than the other sites in aggregates 1 – 2
253 mm, while in young forest P-PO₃³⁻ was significantly higher for 0.5 – 1 mm in comparison to
254 grassland and cropland.

255 Significant differences in physico-chemical parameters between aggregates size classes
256 were found, mainly at the young forest and forest site, and between the classes < 0.5 mm and
257 the other classes. The aggregates size classes < 0.5 mm at the young forest and forest sites had
258 significantly lower SOC concentrations than bulk soil and most larger size classes, while their
259 C/N was higher (Fig. S3). Similarly, the water content of < 0.25 mm was significantly lower than
260 most aggregates sizes at young forest, forest and grassland sites. In contrast, soil aggregates <
261 0.5 mm at grassland showed significantly higher N-NO₃⁻ concentrations than other soil
262 aggregate sizes or bulk soil (Fig. S4). The sand content was higher in cropland and lower in
263 grassland and was higher in aggregate size classes < 0.5 mm regardless of the land use (Fig. S5).
264 In contrast, the silt content was lower in cropland and higher in grassland, while clay content
265 was lower in young forest. Both silt and clay contents tend to decrease in aggregate size classes
266 < 0.5 mm. The different fractions of SOM were lower in cropland and higher in grassland, while
267 labile SOM was higher in aggregate size classes 2 -5 and 1 -2 mm and stable and refractory SOM
268 both tend to decrease in aggregate size classes < 0.5 mm (Fig. S6).

269

270 **3.2. Variation in microbial gene abundance between land uses and soil aggregate size classes**

271 All microbial gene abundances investigated showed significant differences between land
272 use types for at least one soil aggregate size class or bulk soil (Fig. 1, Fig. S7-S9, Table S2). The
273 cropland site consistently (i.e. across bulk soil and soil aggregates) showed lower abundance of
274 bacterial 16S rRNA, *nifH*, *narG*, *nirS* and *nosZ* genes, while *amoA* bacteria (AOB) was lower in
275 grassland (Fig. S8) and *amoA* archaea (AOA) in young forest (Fig. 1, S8). In contrast, the forest
276 site tends to harbour the highest abundance for the different aggregate sizes of bacterial and
277 archaeal 16S rRNA, AOB and AOA genes (Fig. S7, S8), while the *nifH*, *narG* and *nirS* genes
278 showed the highest abundance in young forest site (Fig. 1, S8, S9), and *nosZ* gene in grassland
279 site (Fig. 1, S9).

280 Significant effects of aggregate size within individual land uses were found (one-way
281 ANOVA and Tukey HSD) for all microbial amplicon abundances investigated, except archaeal
282 16S rRNA, fungal ITS, and AOA (Fig. S7-S9). However, significant pairwise differences were only
283 found for the young forest (for bacterial 16S rRNA, *nifH*, and *narG* genes) and forest sites (for
284 AOB, *narG*, *nirS* and *nosZ* genes). Trends at the young forest site were similar, where genes
285 abundances were overall found relatively high in 0.5 -1.0 mm aggregates and relatively low in
286 2.0-5.0 mm and < 0.25 mm aggregates (Fig. 2). For the forest site a similar trend is also found,
287 the abundances being higher in the 0.25 – 0.5 and 0.5 – 1.0 mm aggregates than in the other
288 aggregate size fractions (Fig. 2).

289
290 **3.3. Changes in potential greenhouse gas fluxes between land uses and soil aggregate size**
291 **classes**

292 The types of land use and moisture levels were the main factors differentiating GHG
293 fluxes, although differences between land uses were not as strong as for microbial abundances
294 and consistent across land uses. Greenhouse gas fluxes were significantly different between
295 land use types at both moisture levels for at least one soil aggregate size, except for NO at field
296 moisture (Fig. S10, S11). The CO₂ emissions were significantly different (Tukey HSD) only for

297 0.5 – 1 mm and bulk soil between cropland and forest site, and also between grassland with
298 cropland and young forest sites for the bulk soil (Fig. 3, S10). At elevated moisture, CO₂
299 emissions were consistently significantly lower in cropland compared to grassland sites
300 regardless of the aggregates size classes and bulk soil (Fig. 3, S10). Overall, the CO₂ emissions
301 were significantly different between soil moisture levels, and mainly higher at the elevated
302 moisture content than at field moisture content (Fig. S10). The other GHG fluxes showed large
303 standard deviation (Fig. 3) and overall significant differences between land use types for a few
304 specific aggregate size classes such as < 0.25 (CH₄ elevated moisture), 0.25 – 0.5 (NO, N₂O soil
305 moisture), 1.0 – 2.0 (CH₄ both moisture levels and N₂O field moisture), 5.0 – 10.0 mm (CH₄ and
306 N₂O elevated moisture) (Fig. S10, S11).

307 Within the separate land use types, significant effects of aggregate size at field moisture
308 were only observed for CH₄ at the forest site and for NO at the grassland site. The 0.5 – 1.0 mm
309 aggregates acted as a sink for CH₄ at field moisture while the other aggregates classes were
310 sources of CH₄ (Fig. 4). The aggregate size classes < 0.5 mm from grassland were found to be
311 sources of NO, while larger size classes were sinks at field moisture (Fig. 4). At elevated
312 moisture, the bulk soil showed significantly lower CO₂ emissions than the aggregates size
313 classes, while it was a source of CH₄ and aggregates size classes (except 2.0 – 5.0 mm) were
314 sinks (Fig. 4).

315

316 ***3.4. Relationship between microbial gene abundance, potential greenhouse gases and soil*** 317 ***characteristics***

318 When the correlations were performed on all the land uses, bacteria, fungi and *nosZ*
319 gene abundances showed similar and significant positive correlations with the following soil
320 characteristics: labile SOM, stable SOM, refractory SOM, SOC, total N, and silt for all land uses
321 combined (Fig. 5a). The *narG*, *nirS* and *nifH* gene abundances showed significant positive
322 correlations with silt and carbonate contents and P-PO₄³⁻ concentrations (Fig. S2, S4-S5). In
323 contrast, AOB, AOA and archaea gene abundances showed negative correlations with silt and

324 carbonate contents, but positive correlations with soil water content, N-NO³⁻ concentration and
325 sand content (Fig. 5a). The CO₂ emissions at elevated moisture for the combined land uses were
326 strongly and positively correlated ($\rho > 0.5$) with the three SOM pools, total N, SOC, carbonate
327 and silt, but negatively with sand content ($\rho = -0.74$; Fig. 5b). The CO₂ and CH₄ fluxes at field
328 moisture showed significant and positive correlations with the three SOM pools, total N and SOC.
329 The other GHG fluxes showed significant correlations with only a few specific variables (Fig. 5b).
330 Most gene abundances were significantly and positively correlated to CO₂ emissions at elevated
331 moisture, except AOB, archaea and AOA genes which were negatively correlated (see
332 supplementary and Fig. S12 for details).

333 The heatmaps for the separate land uses did not reveal similar patterns across land use
334 types but unique to each land use, even for young forest and forest sites where significant
335 differences in gene abundances between soil aggregate sizes were found (Fig. 6, S13, S14).
336 Hence, at the young forest site, the N contents and to a lesser extent SOM contents (especially
337 the labile SOM pool) were positively correlated to bacteria, *nifH*, AOB, *narG* and *nirS* genes (Fig.
338 6). At the forest site, different parameters explained the differences in genes abundance
339 between soil aggregate sizes; soil texture explained the distribution of several gene abundances,
340 with clay content positively correlated with *nifH*, bacteria, *narG* and AOB genes and sand with
341 fungi, while sand content was negatively correlated with *nosZ*, and *nirS* genes.

342 The correlations between GHG fluxes and soil properties showed no similar patterns
343 across land uses and relatively low number of correlations (Fig. S13). At the grassland site,
344 where most differences in GHG fluxes between soil aggregate sizes were found, the CH₄ fluxes at
345 field moisture were positively correlated to labile, stable and refractory SOM content, but
346 negatively correlated to these SOM fractions at elevated moisture (Fig. S13). The correlations
347 between gene abundances and GHG fluxes for each land use are presented in supplementary
348 material (Fig. S14)

349

350 4 Discussion

351

352 *4.1 Land use is a dominant explaining factor for microbial gene abundance in soil*

353 The type of land use was the main factor of the microbial abundance and the nitrogen
354 cycling community in soils studied. Regardless of the gene investigated, gene abundances were
355 always affected by the different types of land use. The different types of land use and
356 management were previously found to affect the abundance of microorganisms (Enwall et al.,
357 2010; Hallin et al., 2009; Lauber et al., 2008; Leininger et al., 2006; Ma et al., 2008; Morales et al.,
358 2010; Wallenstein and Vilgalys, 2005). This study present a comprehensive evaluation of the
359 distribution of N cycling genes across land uses with similar parent material (fluvial sediments)
360 and climate (co-located sites).

361 Cropping clearly had a negative effect on the abundance of microorganisms in soil and
362 most of their N functions. The SOC and total N concentrations explained the distribution of
363 bacteria, fungi and *nosZ* gene, highlighting that the depletion of SOC and total N in cropland (Fig.
364 S3) due to soil management (e.g. tillage), soil erosion and plant harvest, limit the abundance of
365 microorganisms. Soil tillage was found to have a direct and negative effect on the biomass of
366 bacteria and fungi (Muruganandam et al., 2009; Helgason et al., 2010), and also on *narG* gene
367 abundance (Chèneby et al., 2009). Hence, the negative effect of cropping on microbial
368 communities is likely due to a combination of factors limiting microbial growth. In contrast, the
369 AOA and AOB were abundant in cropland, likely due to application of fertiliser (containing NH_4)
370 that maintains AOA and AOB and stimulates nitrification which was supported by the significant
371 correlations of the ammonium oxidizing microorganisms with NO_3^- concentration and soil water
372 content. However, distinct drivers of each community were also found across land uses, such as
373 SOC/N and sand content for AOB, and total N, thermally more stable SOM and clay contents for
374 AOA (Fig. 5a). Thus, it further supports the idea that despite AOA and AOB delivering the same
375 function, the two communities live in different niches/microhabitats with specific environments
376 stimulating their activity separately (Prosser and Nicol, 2008). Low soil pH and low NH_4^+

377 concentration were found to be important conditions favouring *amoA* archaea abundance while
378 the contrary was found for *amoA* bacteria (Leininger et al., 2006; Verhamme et al., 2011).
379 However, in the current study the soil pH was above 7 and both bacterial and archaeal *amoA*
380 showed strong positive correlation with NO_3^- and NH_4^+ for archaea, showing that these drivers
381 are not the only ones responsible for niche differentiation of *amoA*. Hence, the quantity and
382 quality of SOM might play an important role in the studied soil, as organic C can differently
383 inhibit or stimulate ammonia oxidizer (Erguder et al., 2009).

384 The community showing the highest abundance in young forest (i.e. *nifH*, *narG* and *nirS*
385 genes) showed a strong and positive correlation to phosphate concentration which was higher
386 in the young forest and could be a limiting factor in the other land use (Table 1, Fig. S3). Their
387 high abundance could also be related to the location of the site, with a slightly different soil type
388 (Epigleyic Fluvisol for young forest and Mollic Fluvisols for the other sites) which is also
389 younger (70 yr against 250-350 yr). Furthermore, the site is located along the Danube River,
390 subjected to flood (~ 10 days yr^{-1}), creating anaerobic conditions over long period of time that
391 would favour the denitrification and N fixation processes. In contrast, the other sites are
392 protected from flood by a dike. The *nifH* gene abundance was found to be higher in forest soil
393 than in agricultural soil (Morales et al., 2010). In contrast, for the communities with higher
394 abundance at the forest site (i.e. bacterial and archaeal 16S rRNA genes, AOB and AOA),
395 different variables were correlated, without a common variable explaining microbial
396 distribution. Hence, this result highlights the complexity of the variables explaining microbial
397 distribution in forest soil (Levy-Booth et al., 2014). The fungal ITS and *nosZ* genes showed
398 similar factors explaining their distribution (i.e. SOC, N, SOM and NO_3^-). Fungi in soils were
399 found to produce N_2O , which in return could be reduced into N_2 by bacteria, which could explain
400 the similar factors between fungal ITS and *nosZ* gene (Maeda et al., 2015). Furthermore, *nosZ*
401 gene distribution showed different factors than *narG* and *nirS* genes, suggesting that the
402 different steps of the denitrification do not simultaneously occur within the same microhabitat
403 which is expected due to the existence of *nosZ* in bacteria lacking other genes for denitrification

404 and the different environment required to perform the different steps of denitrification. Thus,
405 there is a niche differentiation of the different steps of the denitrification, with SOM quantity
406 and quality (directly related to the plant residues input and root exudates) playing a key role for
407 *nosZ* gene abundance, while *narG* and *nirS* genes were both regulated by the P, carbonate and
408 silt concentration.

409

410 ***4.2 Soil aggregate size is explaining minor factor for microbial gene abundance in soil***

411 Soil aggregate size was a minor factor in explaining nitrogen genes abundance,
412 compared to land use. The effects of soil aggregate size classes on gene abundances was specific
413 to the land use type and not present for all genes or land uses studied. Neuman et al. (2013)
414 found that the size of soil aggregates was the dominant factor in the abundance of bacterial,
415 archaeal and fungal community, over soil management (i.e. fertilisation). However, they
416 investigated microaggregates (0.002 – 0.020 mm, 0.020 – 0.063 mm, > 0.063 mm) and the silt
417 and clay fractions (< 0.002 mm), which could physically protect organisms against
418 environmental changes. Hence, the current study shows that the sizes of macroaggregates are
419 not the main factor determining microbial distribution and N functional guilds after land use
420 type, whereas aggregates < 0.063 mm could have a greater effect on the distribution of
421 microbial communities.

422 The presence or absence of differences in gene abundance between soil aggregates in
423 different land use may be related to the balance between stability and instability of the
424 microhabitats, hindering or promoting differentiation of specific microhabitats and associated
425 microbial communities. The low variation in gene abundance for cropland and grassland may be
426 related to the soil aggregates and organic matter turnover, which is expected to be higher due to
427 anthropogenic activity such as tillage and plant harvest (Blaud et al., 2014; Six et al., 2002, 2000;
428 Tisdall and Oades, 1982). The lower variation in microbial abundance between soil aggregate
429 size fractions in grassland in comparison to young forest and forest, might be explained by a
430 high organic matter input due to fine grass root system and root exudates, resulting in the

431 highest SOC and total N concentration in comparison to the other land uses, and no significant
432 difference in their concentrations between grassland aggregate sizes classes (Fig. S3).
433 Furthermore, forest sites were likely to show a more stable temperature and soil moisture
434 regime throughout the year than cropland and grassland because of the tree cover, as well as a
435 different quantity and quality of plant input that affected SOM concentration in soil aggregate
436 size classes (Fig. S6). Overall, specific drivers for each land use are responsible of the
437 distribution of gene abundance in soil aggregates, such as total N and labile SOM that explained
438 bacteria, *nifH*, AOB, *narG* and *nirS* genes distribution for young forest, while soil texture,
439 especially clay content, was explaining most gene distribution in forest. In contrast, for cropland
440 and grassland organic C and silt content respectively, explained few genes distribution.

441 At the forest and young forest sites, the size of soil aggregates was an important factor in
442 the abundance of several microbial communities and functional genes, with specific sizes
443 harbouring higher gene abundances. Furthermore, a similar pattern of distribution was found
444 between functional genes at a specific site, suggesting that these functions coexist in similar
445 niches. Hence, the aggregate size class 0.5 – 1.0 mm consistently showed the highest gene
446 abundance regardless of the specific microbial functions, possibly hosting a high number of
447 active microbial functions, and is within the range of soil aggregates that characterise fertile
448 soils as described by Shein (2005). However, some dissimilarities were present, such as the soil
449 aggregate size class 1.0 – 2.0 mm which showed high gene abundances at the young forest while
450 low gene abundances were found at the forest site. Thus, differences between similar land use,
451 such as tree cover, and soil characteristics may also play a role in gene abundance distribution
452 within soil aggregate size classes. Although those genes preferentially colonised similar niches,
453 which differ in their distribution across land uses, different factors were responsible for their
454 abundances in the young forest and forest site.

455

456 ***4.3 Effects of land use and soil aggregate size on potential greenhouse gas fluxes***

457 The potential GHG fluxes were affected by land use, soil moisture levels and to a lesser
458 extent soil aggregate size, but the effects were far less pronounced than for microbial gene
459 abundance, and inconsistent across land use and soil aggregates. This was partly due to the high
460 variability in the measure of GHG fluxes, but also revealed differences compared to the
461 microbial gene abundance. Hence, the effect of land use on the bulk soil samples were mainly
462 found for CO₂ emissions, while for the other GHG only specific soil aggregate sizes revealed the
463 potential effect of land use. The different effect of land use found on GHG fluxes between soil
464 aggregate size classes compared to the bulk soil may be linked to different porosity present for
465 each size and how land use affects it differentially (Rabbi et al., 2016). Thus, working on bulk
466 soil may mask some potential GHG fluxes (Kravchenko et al., 2014). However, it should be
467 acknowledged that each soil aggregate size was in artificial conditions for the GHG
468 measurement (e.g. air fluxes), likely leading to different behaviour than *in situ*. The CO₂
469 emissions were consistently lower in cropland compared to the other sites regardless of the soil
470 water content, indicating the potential low microbial activity in cropland due to SOM depletion
471 also supported by the low bacterial gene abundance, but also strong correlations with most
472 genes abundance. The other GHG fluxes showed inconsistent effect of land use depending on
473 soil moisture and soil aggregate size, highlighting the complexity of drivers of CH₄, NO and N₂O
474 fluxes. Only few correlations were found between CH₄, NO and N₂O fluxes and genes abundance,
475 showing the difficulty to relate gene abundance and GHG fluxes, due to the high variability of
476 GHG fluxes and possible dissimilarity between genes and activity.

477 Change in soil moisture had significant effects on GHG fluxes, although it varies between
478 GHG, land use, and soil aggregate size classes. Higher CO₂ emissions were consistently found at
479 elevated soil moisture compared to field moisture across all land use, highlighting the
480 importance of soil moisture for microbial activity and CO₂ emissions (Sey et al., 2008). For CH₄,
481 NO and N₂O the effect of increased soil moisture was not as consistent as for CO₂, indicating that
482 other factors limit their fluxes. Surprisingly, increasing soil water content in the current study

483 did not necessarily increase the CH₄ production, as might be expected because methanogens are
484 more active in high water content/anaerobic soils. The CH₄ was either emitted or consumed
485 depending on the soil water content for a specific land use and soil aggregate size class. This
486 indicates that both methanogens and methane-oxidizing bacteria were present in the same soil
487 aggregates as previously found by Sey et al. (2008) and can co-exist in the same niche. Similarly,
488 increasing soil water content did not increase the anaerobic process of denitrification
489 responsible for NO and N₂O fluxes, indicating that other factors are regulating these fluxes and
490 the microorganisms responsible, or the increase in soil water content was not enough to reach
491 anaerobic conditions.

492 Overall, the GHG fluxes did not occur in a specific aggregate size class within a land use
493 as found for microbial gene abundances in forest sites. Previous studies found higher CO₂
494 emissions in microaggregates whilst acting as sinks of CH₄ (Sey et al., 2008). However, CO₂
495 emissions were also shown to be highly sensitive to water filled pore space (WFPS), with no
496 difference in emissions between aggregate size at 60% WFPS; microaggregates acted as sinks of
497 CH₄ at 20% WFPS but a source at higher WFPS (Ramakrishnan et al., 2000; Sey et al., 2008).
498 However, in the current study, elevated soil moisture did not reveal more significant differences
499 than at soil moisture in GHG fluxes between soil aggregates, indicating that other factors may
500 drive differences or that the size of soil aggregate may not be an important driver for GHG fluxes.

501

502 **5. Conclusions**

503 This study demonstrates that land use is the main factor in explaining abundance of
504 nitrogen genes and greenhouse gas fluxes, while soil aggregate size class was a minor factor.
505 This goes against our initial hypothesis suggesting that different microbial functions are
506 preferentially hosted or fostered by specific size of aggregates. This is due to the stronger
507 difference in soil physico-chemical characteristics between land use types than between soil
508 aggregate sizes. Cropping had a clear negative effect on the abundance of most microbial
509 communities, likely due to the depletion of SOC and total N by tillage, plant harvest, and soil

510 erosion. Although soil aggregate size was not a dominant factor, it affected the distribution of
511 the N functional communities at the semi-natural forest sites, showing that some microbial
512 functions are probably related to specific microhabitats (i.e. the architecture and distribution of
513 pores filled with water and air, the availability of organic matter and other nutrients) in soil,
514 where anthropogenic activity is limited, allowing differences between microhabitats to develop.
515 However, no specific size of soil aggregates enhanced the abundance of any specific microbial
516 function across all four land uses. Soil aggregate size had little effect on GHG fluxes, indicating
517 that the size of soil aggregates may not have much effect on GHG fluxes but it also highlights the
518 difficulties of measuring GHG fluxes in aggregates.

519 This study only addresses a single point in time, limiting our understanding of the
520 distribution of microbial functions over soil aggregates of different size. Further studies are
521 needed, taking into consideration the dynamics of soil aggregates and its relation with microbial
522 communities by sampling at multiple time points, work on a wider range of aggregate size
523 classes (e.g. size classes < 0.25 mm) and land use types. Furthermore, combining microbiology
524 and soil architecture (e.g. x-ray tomography) as well as nutrient availability in local and time
525 scale, would fully reveal the physical distribution of microhabitats, the microbial communities
526 and functions among soil aggregates. Comparing microbial functions between soil aggregates of
527 varying size from a specific land use (e.g. forest) but from different locations or soil types may
528 also provide more insight into the role of soil aggregates in microbial functioning.

529

530 **Acknowledgements**

531 This work was supported by the European Commission 7th Framework Program as a Large
532 Integrating Project, SoilTrEC (www.soiltrec.eu), Grant Agreement No. 244118.

533

534 **References**

535 Amézketa, E., 1999. Soil aggregate stability: a review. *J. Sustain. Agr.* 14, 83–151.

536 An, S., Mentler, A., Mayer, H., Blum, W.E.H., 2010. Soil aggregation, aggregate stability, organic
537 carbon and nitrogen in different soil aggregate fractions under forest and shrub
538 vegetation on the Loess Plateau, China. *CATENA* 81, 226–233.

539 Ball, B.C., 2013. Soil structure and greenhouse gas emissions: a synthesis of 20 years of
540 experimentation. *Eur. J. Soil Sci.* 64, 357–373.

541 Banwart, S., 2011. Save our soils. *Nature* 474, 151–152.

542 Barros, N., Salgado, J., Feijóo, S., 2007. Calorimetry and soil. *Thermochimica Acta*, XIVth ISBC
543 Proceedings Special Issue Fourteenth conference of the International Society for
544 Biological Calorimetry 458, 11–17.

545 Beauchamp, E.G., Seech, A.G., 1990. Denitrification with different sizes of soil aggregates
546 obtained from dry-sieving and from sieving with water. *Biol. Fertil. Soils* 10, 188–193.

547 Blaud, A., Lerch, T.Z., Chevallier, T., Nunan, N., Chenu, C., Brauman, A., 2012. Dynamics of
548 bacterial communities in relation to soil aggregate formation during the decomposition
549 of ¹³C-labelled rice straw. *Appl. Soil Ecol.* 53, 1–9.

550 Blaud, A., Chevallier, T., Virto, I., Pablo, A.-L., Chenu, C., Brauman, A., 2014. Bacterial community
551 structure in soil microaggregates and on particulate organic matter fractions located
552 outside or inside soil macroaggregates. *Pedobiologia* 57, 191–194.

553 Blaud, A., Menon, M., van der Zaan, B., Lair, G.J., Banwart, S.A., 2017. Chapter Five - Effects of Dry
554 and Wet Sieving of Soil on Identification and Interpretation of Microbial Community
555 Composition, in: Sparks, S.A.B. and D.L. (Ed.), *Advances in Agronomy, Quantifying and
556 Managing Soil Functions in Earth's Critical Zone Combining Experimentation and
557 Mathematical Modelling*. Academic Press, pp. 119–142.

558 Bossuyt, H., Six, J., Hendrix, P.F., 2002. Aggregate-protected carbon in no-tillage and
559 conventional tillage agroecosystems using carbon-14 labeled plant residue. *Soil Sci. Soc.
560 Am. J.* 66, 1965–1973.

561 Bronick, C.J., Lal, R., 2005. Soil structure and management: a review. *Geoderma* 124, 3–22.

562 Chèneby, D., Brauman, A., Rabary, B., Philippot, L., 2009. Differential responses of nitrate
563 reducer community size, structure, and activity to tillage systems. *Appl. Environ.*
564 *Microbiol.* 75, 3180–3186.

565 Chotte, J.L., Schwartzmann, A., Bally, R., Jocteur Monrozier, L., 2002. Changes in bacterial
566 communities and *Azospirillum* diversity in soil fractions of a tropical soil under 3 or 19
567 years of natural fallow. *Soil Biol. Biochem.* 34, 1083–1092.

568 Davinic, M., Fultz, L.M., Acosta-Martinez, V., Calderón, F.J., Cox, S.B., Dowd, S.E., Allen, V.G., Zak,
569 J.C., Moore-Kucera, J., 2012. Pyrosequencing and mid-infrared spectroscopy reveal
570 distinct aggregate stratification of soil bacterial communities and organic matter
571 composition. *Soil Biol. Biochem.* 46, 63–72.

572 Enwall, K., Throbäck, I.N., Stenberg, M., Söderström, M., Hallin, S., 2010. Soil resources influence
573 spatial patterns of denitrifying communities at scales compatible with land management.
574 *Environ. Microbiol. Appl. Environ. Microbiol.* 76, 2243–2250.

Erguder, T.H., Boon, N., Wittebolle, L., Marzorati, M., Verstraete, W., 2009. Environmental factors
shaping the ecological niches of ammonia-oxidizing archaea. *FEMS Microbiol. Rev.* 33,
855–869.

575 Hallin, S., Jones, C.M., Schloter, M., Philippot, L., 2009. Relationship between N-cycling
576 communities and ecosystem functioning in a 50-year-old fertilization experiment. *ISME J.*
577 3, 597–605.

578 Hayden, H.L., Drake, J., Imhof, M., Oxley, A.P.A., Norng, S., Mele, P.M., 2010. The abundance of
579 nitrogen cycle genes *amoA* and *nifH* depends on land-uses and soil types in South-
580 Eastern Australia. *Soil Biol. Biochem.* 42, 1774–1783.

581 Helgason, B.L., Walley, F.L., Germida, J.J., 2010. No-till soil management increases microbial
582 biomass and alters community profiles in soil aggregates. *Appl. Soil Ecol.* 46, 390–397.

583 IUSS Working Group WRB, 2014. World reference base for soil resources 2006, World soil
584 resources rep 103. ed. FAO, Rome.

585 Izquierdo, J., Nüsslein, K., 2006. Distribution of extensive *nifH* gene diversity across physical soil
586 microenvironments. *Microbial Ecol.* 51, 441–452.

587 Kaiser, M., Kleber, M., Berhe, A.A., 2015. How air-drying and rewetting modify soil organic
588 matter characteristics: An assessment to improve data interpretation and inference. *Soil*
589 *Biol. Biochem.* 80, 324–340.

590 Kravchenko, A.N., Negassa, W.C., Guber, A.K., Hildebrandt, B., Marsh, T.L., Rivers, M.L., 2014.
591 Intra-aggregate pore structure influences phylogenetic composition of bacterial
592 community in macroaggregates. *Soil Sci. Soc. Am. J.* 78, 1924.

593 Lair, G.J., Zehetner, F., Hrachowitz, M., Franz, N., Maringer, F.-J., Gerzabek, M.H., 2009. Dating of
594 soil layers in a young floodplain using iron oxide crystallinity. *Quatern Geochronol* 4,
595 260–266.

596 Lauber, C.L., Strickland, M.S., Bradford, M.A., Fierer, N., 2008. The influence of soil properties on
597 the structure of bacterial and fungal communities across land-use types. *Soil Biol.*
598 *Biochem.* 40, 2407–2415.

599 Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G.W., Prosser, J.I., Schuster, S.C.,
600 Schleper, C., 2006. Archaea predominate among ammonia-oxidizing prokaryotes in soils.
601 *Nature* 442, 806–809.

602 Lensi, R., Clays-Josser, A., Jocteur Monrozier, L., 1995. Denitrifiers and denitrifying activity in
603 size fractions of a mollisol under permanent pasture and continuous cultivation. *Soil*
604 *Biol. Biochem.* 27, 61–69.

605 Levy-Booth, D.J., Prescott, C.E., Grayston, S.J., 2014. Microbial functional genes involved in
606 nitrogen fixation, nitrification and denitrification in forest ecosystems. *Soil Biol.*
607 *Biochem.* 75, 11–25.

608 Ma, W.K., Bedard-Haughn, A., Siciliano, S.D., Farrell, R.E., 2008. Relationship between nitrifier
609 and denitrifier community composition and abundance in predicting nitrous oxide
610 emissions from ephemeral wetland soils. *Soil Biol. Biochem.* 40, 1114–1123.

611 Maeda, K., Spor, A., Edel-Hermann, V., Heraud, C., Breuil, M.-C., Bizouard, F., Toyoda, S., Yoshida,
612 N., Steinberg, C., Philippot, L., 2015. N₂O production, a widespread trait in fungi. *Sci. Rep.*
613 5.

614 Mangalassery, S., Sjögersten, S., Sparkes, D.L., Sturrock, C.J., Mooney, S.J., 2013. The effect of soil
615 aggregate size on pore structure and its consequence on emission of greenhouse gases.
616 *Soil Till. Res.* 132, 39–46.

617 Mendes, I.C., Bottomley, P.J., 1998. Distribution of a population of *Rhizobium leguminosarum* *bv.*
618 *trifolii* among different size classes of soil aggregates. *Appl. Environ. Microbiol.* 64, 970–
619 975.

620 Morales, S.E., Cosart, T., Holben, W.E., 2010. Bacterial gene abundances as indicators of
621 greenhouse gas emission in soils. *ISME J* 4, 799–808.

622 Muruganandam, S., Israel, D.W., Robarge, W.P., 2009. Activities of nitrogen-mineralization
623 enzymes associated with soil aggregate size fractions of three tillage systems. *Soil Sci.*
624 *Soc. Am. J.* 73, 751.

625 Neumann, D., Heuer, A., Hemkemeyer, M., Martens, R., Tebbe, C.C., 2013. Response of microbial
626 communities to long-term fertilization depends on their microhabitat. *FEMS Microbiol.*
627 *Ecol.* 86, 71–84.

628 Pinheiro, E.F.M., Pereira, M.G., Anjos, L.H.C., 2004. Aggregate distribution and soil organic matter
629 under different tillage systems for vegetable crops in a Red Latosol from Brazil. *Soil Till.*
630 *Res.* 77, 79–84.

631 Poly, F., Ranjard, L., Nazaret, S., Gourbiere, F., Jocteur Monrozier, L., 2001. Comparison of *nifH*
632 gene pools in soils and soil microenvironments with contrasting properties. *Environ.*
633 *Microbiol. Appl. Environ. Microbiol.* 67, 2255–2262.

634 Prosser, J.I., Nicol, G.W., 2008. Relative contributions of archaea and bacteria to aerobic
635 ammonia oxidation in the environment. *Environ. Microbiol.* 10, 2931–2941.

636 R Development Core Team, 2015. R: a language and environment for statistical computing.

637 Rabbi, S.M.F., Daniel, H., Lockwood, P.V., Macdonald, C., Pereg, L., Tighe, M., Wilson, B.R., Young,
638 I.M., 2016. Physical soil architectural traits are functionally linked to carbon
639 decomposition and bacterial diversity. *Sci. Rep.* 6, 33012.

640 Ramakrishnan, B., Lueders, T., Conrad, R., Friedrich, M., 2000. Effect of soil aggregate size on
641 methanogenesis and archaeal community structure in anoxic rice field soil. *FEMS*
642 *Microbiol. Ecol.* 32, 261–270.

643 Schaufler, G., Kitzler, B., Schindlbacher, A., Skiba, U., Sutton, M.A., Zechmeister-Boltenstern, S.,
644 2010. Greenhouse gas emissions from European soils under different land use: effects of
645 soil moisture and temperature. *Eur. J. Soil Sci.* 61, 683–696.

646 Schindlbacher, A., Zechmeister-Boltenstern, S., Butterbach-Bahl, K., 2004. Effects of soil
647 moisture and temperature on NO, NO₂, and N₂O emissions from European forest soils. *J.*
648 *Geophys. Res.* 109, 1–12.

649 Sey, B.K., Manceur, A.M., Whalen, J.K., Gregorich, E.G., Rochette, P., 2008. Small-scale
650 heterogeneity in carbon dioxide, nitrous oxide and methane production from aggregates
651 of a cultivated sandy-loam soil. *Soil Biol. Biochem.* 40, 2468–2473.

652 Shein, E.V., 2005. *Kurs fiziki pochv (A Course of Soil Physics)* [in Russian]. Moscow State Univ.
653 Publishing.

654 Six, J., Bossuyt, H., Degryze, S., Denef, K., 2004. A history of research on the link between
655 (micro)aggregates, soil biota, and soil organic matter dynamics. *Soil Till. Res.* 79, 7–31.

656 Six, J., Conant, R.T., Paul, E.A., Paustian, K., 2002. Stabilization mechanisms of soil organic matter:
657 Implications for C-saturation of soils. *Plant Soil* 241, 155–176.

658 Six, J., Elliott, E.T., Paustian, K., 2000. Soil macroaggregate turnover and microaggregate
659 formation: a mechanism for C sequestration under no-tillage agriculture. *Soil Biol.*
660 *Biochem.* 32, 2099–2103.

661 Tisdall, J.M., Oades, J.M., 1982. Organic matter and water-stable aggregates in soils. *Eur. J. Soil Sci.*
662 33, 141–163.

Verhamme, D.T., Prosser, J.I., Nicol, G.W., 2011. Ammonia concentration determines differential growth of ammonia-oxidising archaea and bacteria in soil microcosms. *ISME J.* 5, 1067–1071.

663 Wallenstein, M.D., Vitgaly, R.J., 2005. Quantitative analyses of nitrogen cycling genes in soils.
664 *Pedobiologia* 49, 665–672.

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686 **Table**

687

688 **Table 1.** Soil characteristics and soil aggregate size distribution of bulk soil samples on a dry
 689 mass basis. Mean value \pm one standard deviation ($n = 3$) are shown.

	Cropland	Young forest	Forest	Grassland
Location	48°09'N, 16°41'E	48°07'N, 16°43'E	48°08'N, 16°39'E	48°11'N, 16°44'E
Soil (0-10 cm) age (yr)	< 70	250-350	250-350	250-350
Water content (%)	11.3 \pm 0.26	14.1 \pm 1.11	17.1 \pm 0.69	12.0 \pm 0.26
Soil pH (H ₂ O)	7.7 \pm 0.14	7.5 \pm 0.07	7.4 \pm 0.17	7.4 \pm 0.09
Organic C (%)	2.4 \pm 0.36	3.2 \pm 0.08	3.8 \pm 0.28	5.0 \pm 0.60
Total N (%)	0.13 \pm 0.01	0.17 \pm 0.01	0.25 \pm 0.02	0.33 \pm 0.04
C _{org} /N	18.1 \pm 1.83	18.5 \pm 1.60	15.1 \pm 1.02	15.0 \pm 0.52
N-NH ₄ ⁺ (mg kg ⁻¹)	1.59 \pm 0.29	0.49 \pm 0.01	0.57 \pm 0.03	4.77 \pm 0.98
N-NO ₃ ⁻ (mg kg ⁻¹)	20.3 \pm 3.07	18.6 \pm 4.00	24.3 \pm 3.13	1.5 \pm 0.66
P-PO ₄ ³⁻ (g kg ⁻¹)	0.35 \pm 0.10	1.13 \pm 0.47	0.85 \pm 0.48	0.59 \pm 0.04
CaCO ₃ (%)	19.0 \pm 1.90	20.6 \pm 1.11	20.4 \pm 0.62	21.1 \pm 1.41
Sand, 63-2000 μ m (%)	32.7	20.2	22.5	8.2
Silt, 2-63 μ m (%)	43.8	63.4	51.2	63.0
Clay, < 2 μ m (%)	23.5	16.4	26.3	28.8
Soil texture	loam	silt loam	silt loam	silt loam
> 10 mm	37.3 \pm 9.1	11.3 \pm 1.0	11.9 \pm 4.4	7.9 \pm 2.4
5.0 - 10.0 mm	14.6 \pm 2.4	15.5 \pm 1.1	18.3 \pm 2.7	21.5 \pm 2.0
2.0 - 5.0 mm	20.5 \pm 4.0	26.1 \pm 3.1	31.2 \pm 2.2	37.8 \pm 3.6
1.0 - 2.0 mm	11.8 \pm 2.4	21.8 \pm 4.1	23.1 \pm 8.4	14.5 \pm 0.5
0.5 - 1.0 mm	6.4 \pm 3.5	9.3 \pm 2.8	5.9 \pm 1.7	5.2 \pm 0.4
0.25 - 0.5 mm	7.1 \pm 4.6	12.7 \pm 2.6	7.5 \pm 2.7	6.9 \pm 0.1
< 0.25 mm	1.9 \pm 1.3	3.3 \pm 0.4	2.0 \pm 0.8	6.1 \pm 0.7

690

691

692

693

694 **Figures captions**

695

696 **Fig. 1** Variation in gene abundance between bulk soil from four land use types. The following
697 genes and microbial communities were targeted: bacterial and archaea (16S rRNA gene), fungi
698 (ITS region), N fixation (*nifH* gene), ammonia oxidizing bacteria and archaea (*amoA* gene,
699 named AOB and AOA, respectively), nitrate reductase (*narG* gene), nitrite reductase (*nirK* gene)
700 and nitrous oxide reductase (*nosZ* gene). All abundances are expressed on the basis of 1 g of dry
701 soil. Mean value \pm one standard deviation ($n = 3$) are shown. Small letters indicate significance
702 ($P < 0.05$) of pairwise differences between land use.

703

704 **Fig. 2.** Variation in gene abundance between bulk soil and six soil aggregates sizes classes from
705 young forest and forest. The following genes and microbial communities were targeted:
706 bacterial and archaea (16S rRNA gene), fungi (ITS region), N fixation (*nifH* gene), ammonia
707 oxidizing bacteria and archaea (*amoA* gene, named AOB and AOA, respectively), nitrate
708 reductase (*narG* gene), nitrite reductase (*nirK* gene) and nitrous oxide reductase (*nosZ* gene).
709 All abundances are expressed on the basis of 1 g of dry mass of the bulk soil or the specific
710 aggregate size fraction. Mean value \pm one standard deviation ($n = 3$) are shown. Small letters
711 indicate significance ($P < 0.05$) of pairwise differences between soil aggregate size classes
712 within a specific land use.

713

714 **Fig. 3.** Variation in GHG fluxes ($\mu\text{g kg}^{-1} \text{h}^{-1}$) between bulk soil from four land use types at field
715 moisture or elevated moisture (40 – 60 % of field capacity). Mean value \pm one standard
716 deviation ($n = 3$) are shown. Small letters indicate significance ($P < 0.05$) of pairwise differences
717 between soil aggregate size classes within a specific land use.

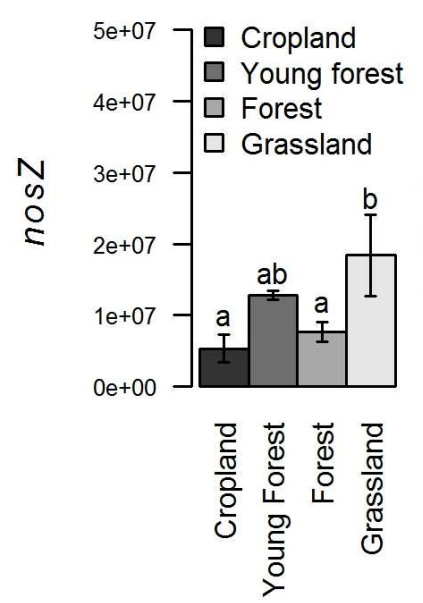
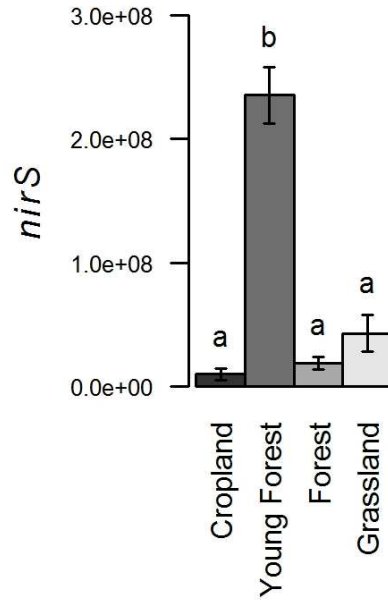
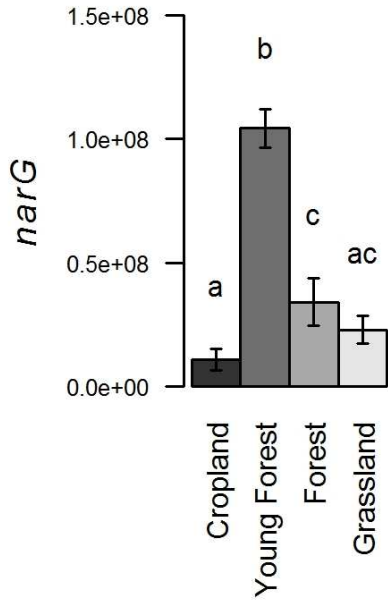
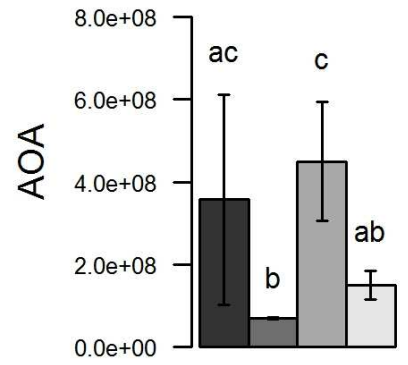
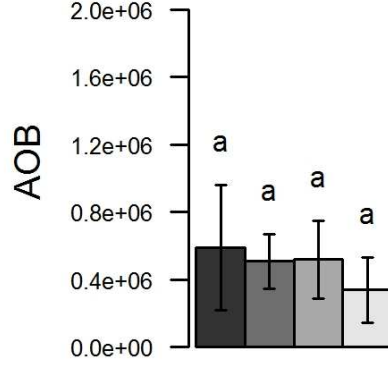
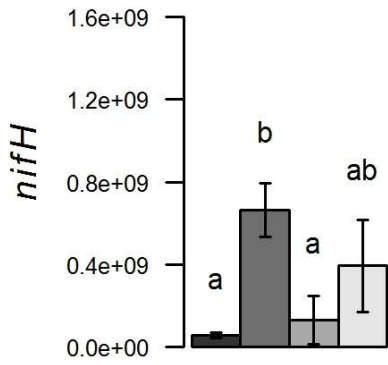
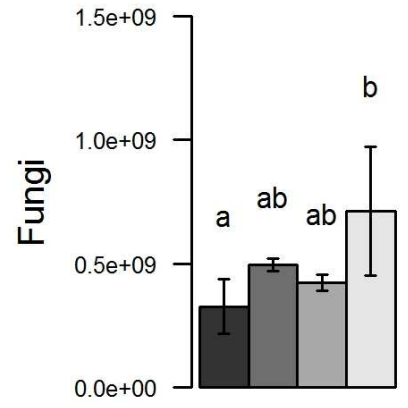
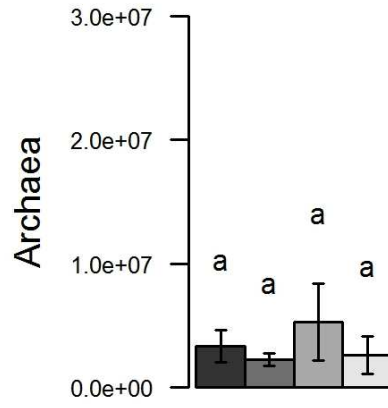
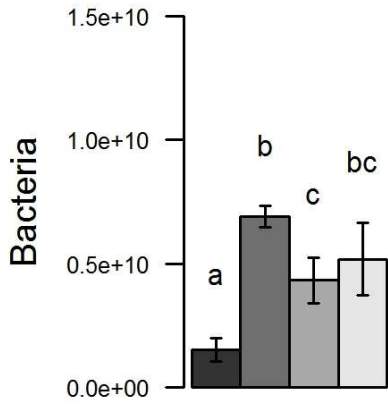
718

719 **Fig. 4.** Variation in GHG fluxes ($\mu\text{g kg}^{-1} \text{h}^{-1}$) between bulk soil and six soil aggregates sizes classes
720 from grassland or forest at field moisture or elevated moisture (40 – 60 % of field capacity).

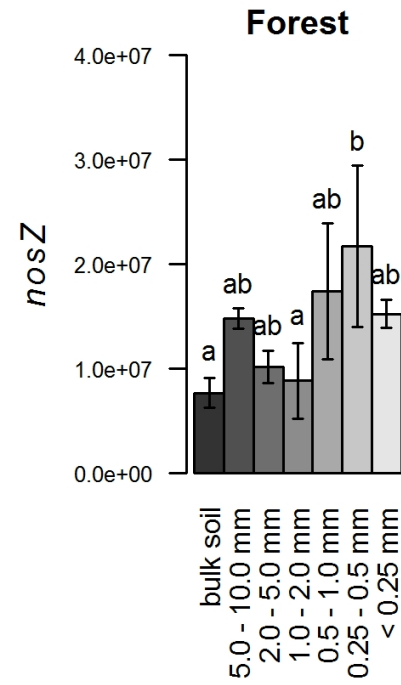
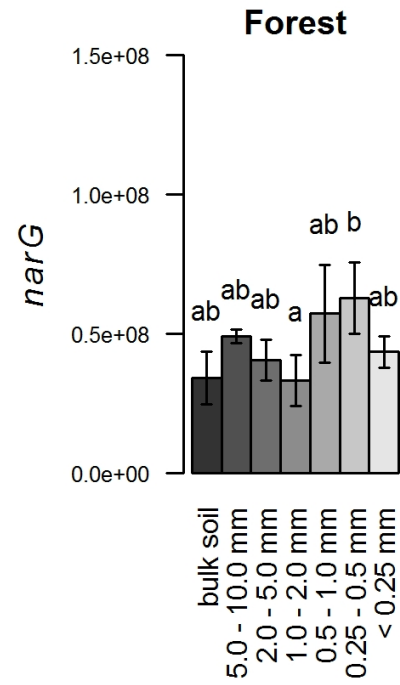
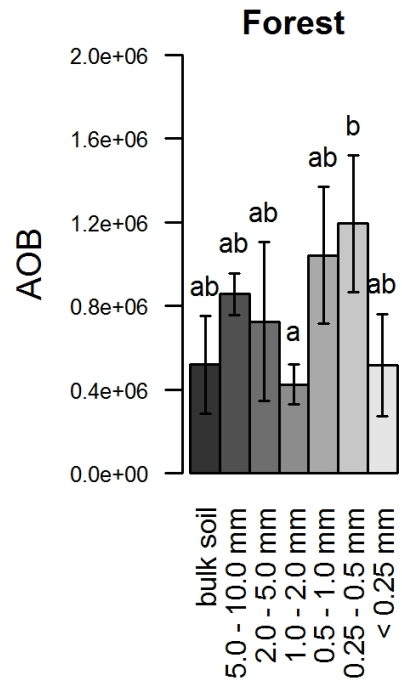
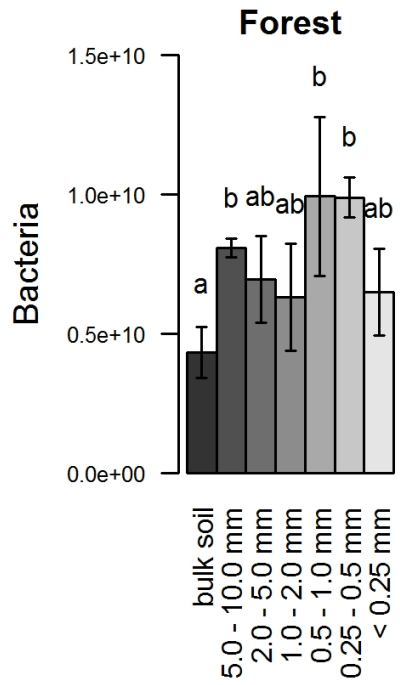
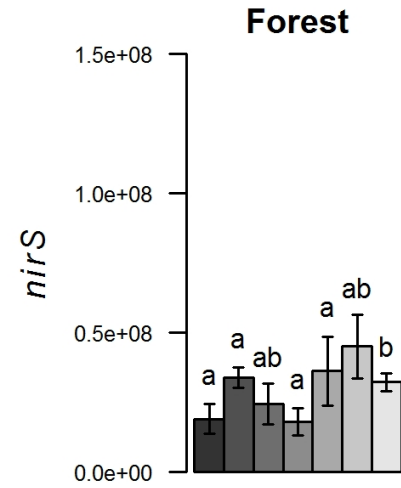
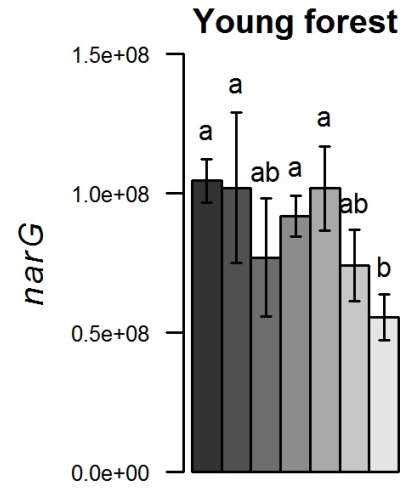
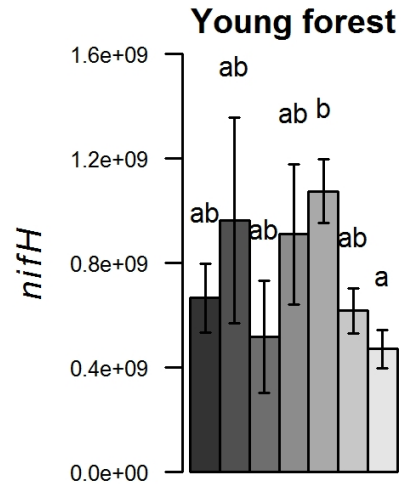
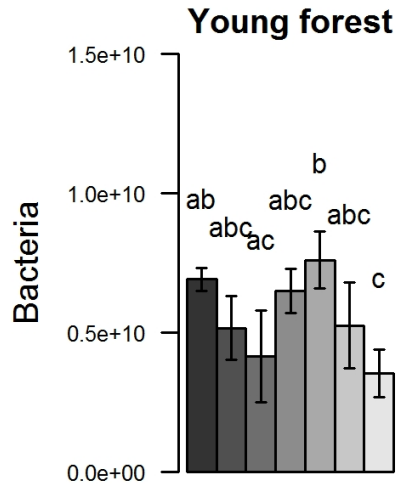
721 Mean value \pm one standard deviation ($n = 3$) are shown. Small letters indicate significance ($P <$
722 0.05) of pairwise differences between soil aggregate size classes within a specific land use.

723
724 **Fig. 5.** Heatmaps of Spearman's rank correlation coefficients ρ between a) soil properties and
725 microbial genes abundance, b) soil properties and greenhouse gas fluxes from samples across
726 six soil aggregates sizes classes (< 0.25, 0.25 – 0.5, 0.5 – 1.0, 1.0 – 2.0, 2.0 – 5.0 and 5.0 – 10.0
727 mm) and four land uses. AOB: *amoA* bacteria; AOA: *amoA* archaea. The ρ values > 0.24 and $< -$
728 0.24 are significant ($P < 0.05$).

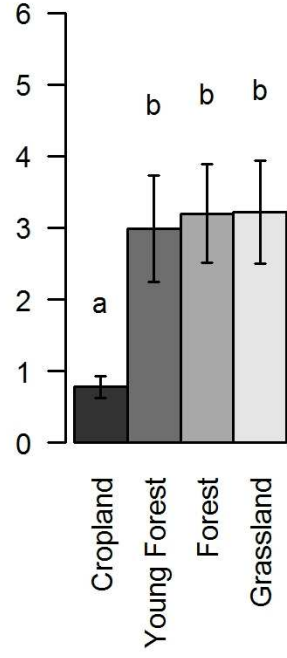
729
730 **Fig. 6** Heatmaps of Spearman's rank correlation coefficients ρ between soil properties and
731 microbial genes abundance from samples across six soil aggregates sizes classes (< 0.25, 0.25 –
732 0.5, 0.5 – 1.0, 1.0 – 2.0, 2.0 – 5.0 and 5.0 – 10.0 mm) and for a) young forest and b) forest sites
733 separately, which showed significant variation in gene abundance with aggregates size classes
734 (refers to figure S13 for the other land uses). AOB: *amoA* bacteria; AOA: *amoA* archaea. The ρ
735 values > 0.47 and < -0.47 are significant ($P < 0.05$).



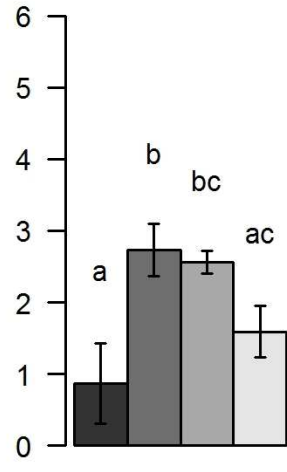
■ Cropland
 ■ Young forest
 ■ Forest
 □ Grassland



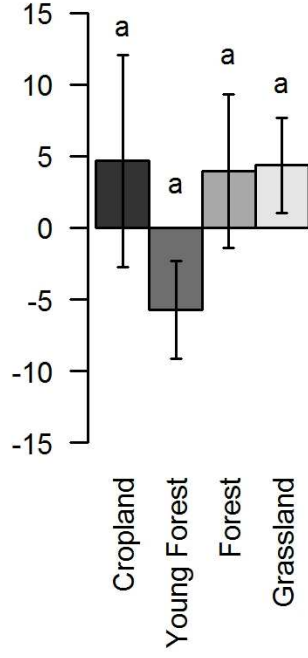
CO₂ elevated moisture ($\mu\text{g kg}^{-1} \text{h}^{-1}$)



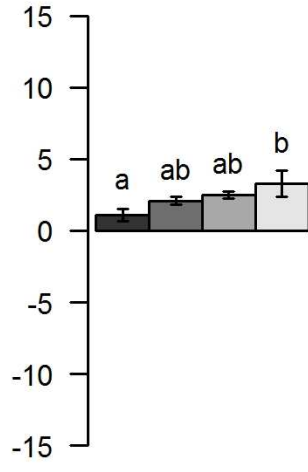
CO₂ field moisture ($\mu\text{g kg}^{-1} \text{h}^{-1}$)



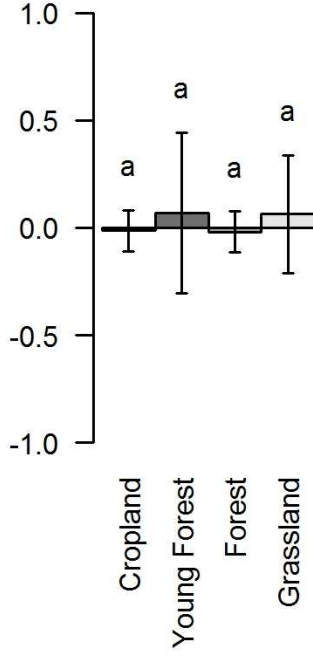
CH₄ elevated moisture ($\mu\text{g kg}^{-1} \text{h}^{-1}$)



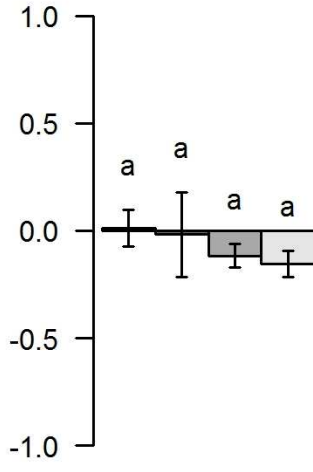
CH₄ field moisture ($\mu\text{g kg}^{-1} \text{h}^{-1}$)



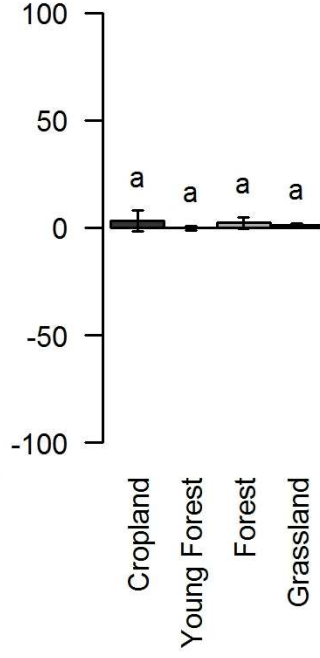
NO elevated moisture ($\mu\text{g kg}^{-1} \text{h}^{-1}$)



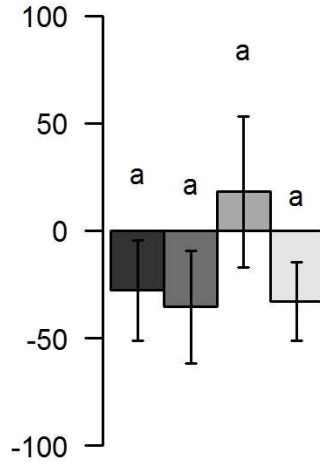
NO field moisture ($\mu\text{g kg}^{-1} \text{h}^{-1}$)

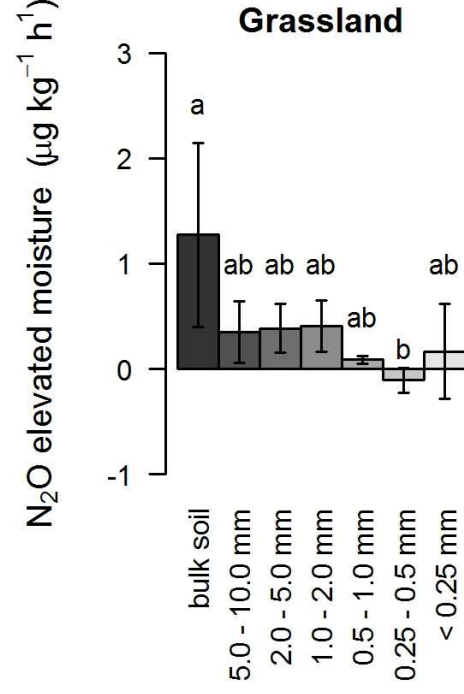
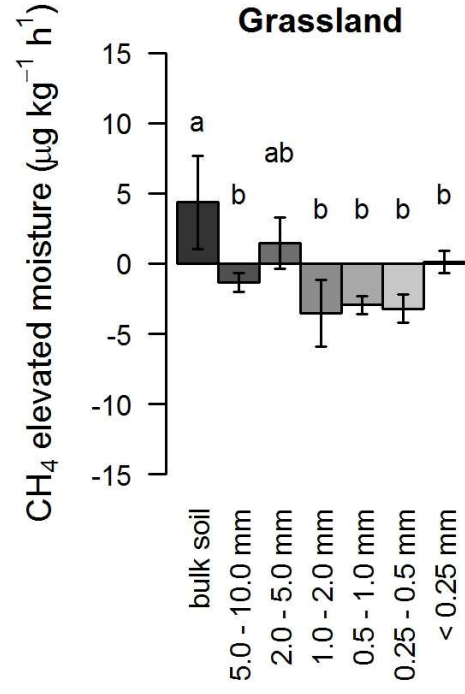
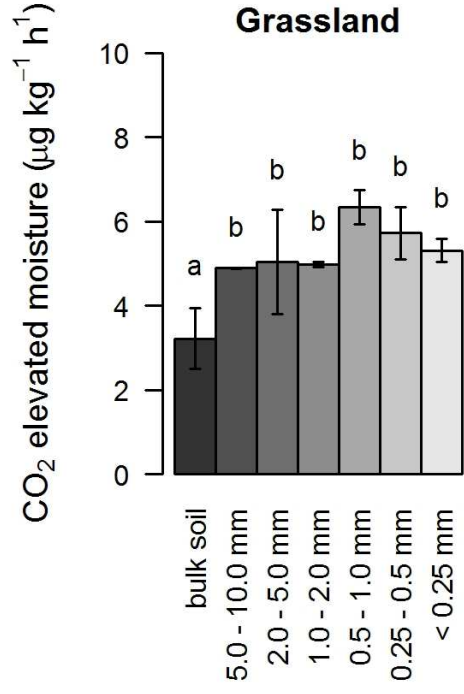
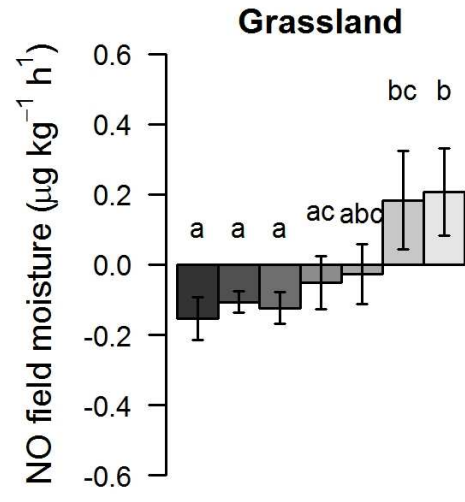
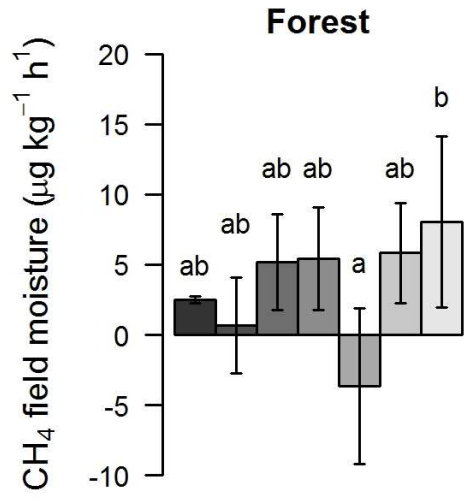


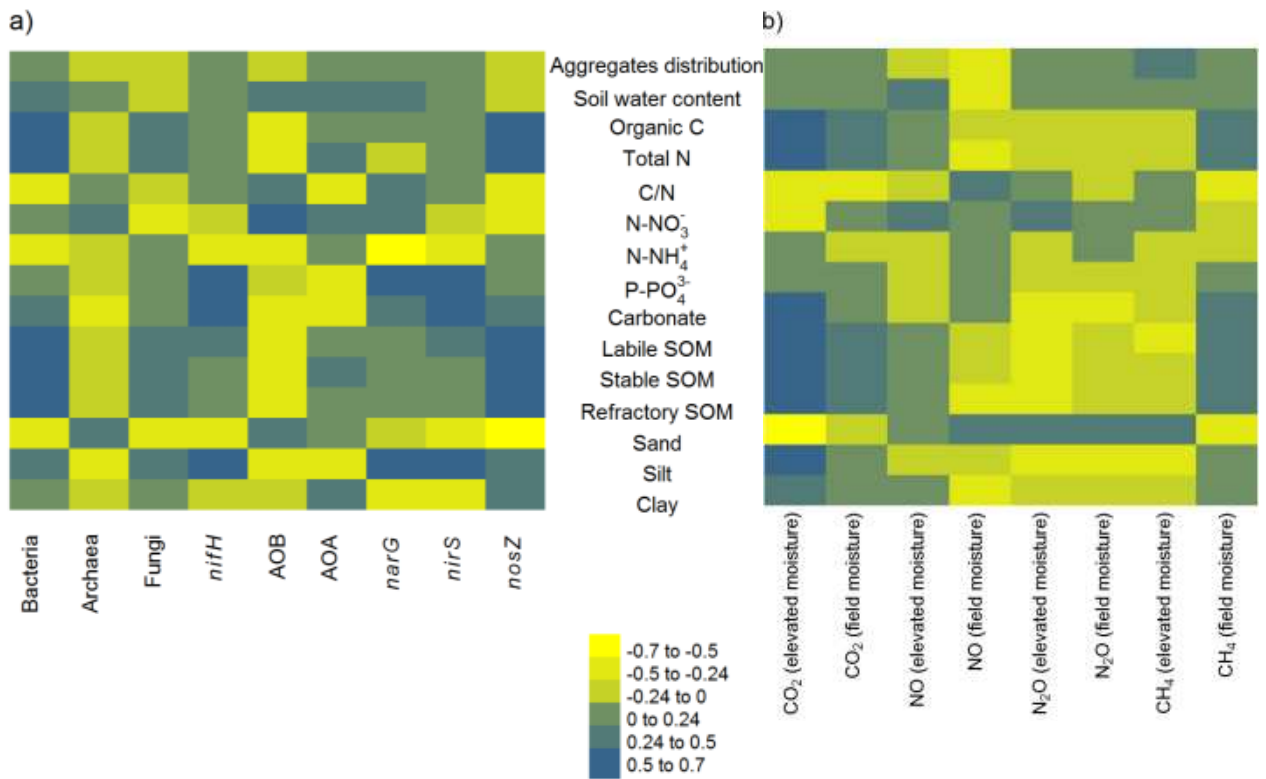
N₂O elevated moisture ($\mu\text{g kg}^{-1} \text{h}^{-1}$)

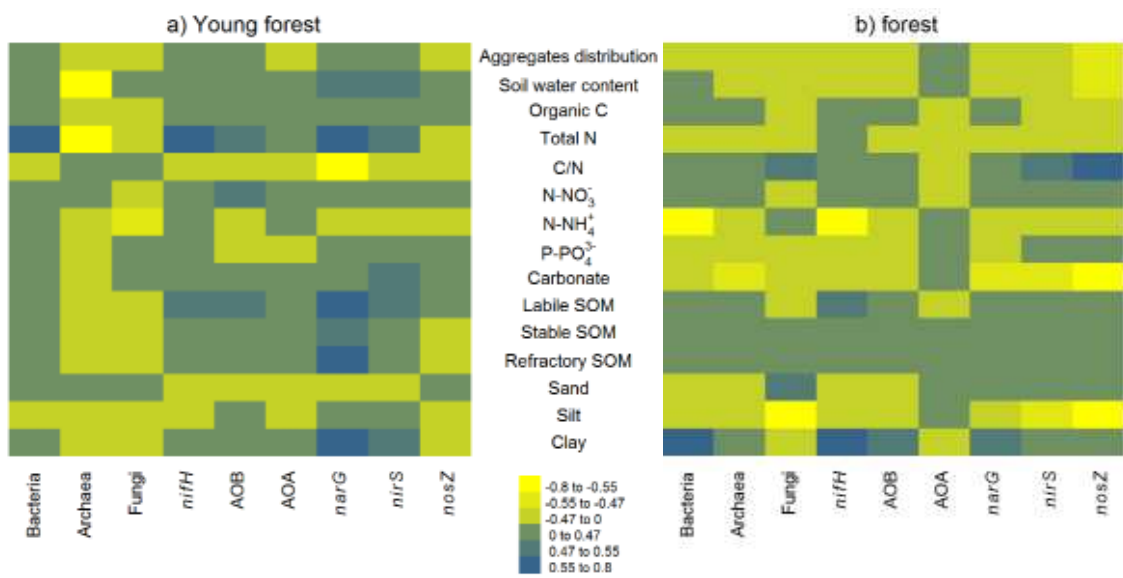


N₂O field moisture ($\mu\text{g kg}^{-1} \text{h}^{-1}$)









The abundance of nitrogen cycle genes and potential greenhouse gas fluxes depends on land use type and little on soil aggregate size

Aimeric Blaud^{a, 1*}, Bas van der Zaan^b, Manoj Menon^{a, 2}, Georg J. Lair^{c, d}, Dayi Zhang^{a, 3}, Petra Huber^c, Jasmin Schiefer^c, Winfried E.H. Blum^c, Barbara Kitzler^e, Wei E.Huang^{a, 4}, Pauline van Gaans^b, and Steve Banwart^a

^a Department of Civil and Structural Engineering, Kroto Research Institute, University of Sheffield, Broad Lane, Sheffield S3 7HQ, United Kingdom.

^b Deltares, Subsurface and Groundwater Systems, Princetonlaan 6-8, 3508 AI Utrecht, the Netherlands.

^c University of Natural Resources and Life Sciences (BOKU), Institute of Soil Research, Vienna, Peter-Jordan-Str. 82, 1190 Vienna, Austria.

^d University of Innsbruck, Institute of Ecology, Sternwartestr. 15, 6020 Innsbruck, Austria.

^e Department of Forest Ecology and Soil, Soil Ecology, Federal Research Centre for Forests, Seckendorff-Gudent-Weg 8, 1131 Vienna Austria.

*Corresponding Author.

E-mail address: aimerick.blaud@gmail.com

¹ Current address: Sustainable Agriculture Science Department, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK.

² Current address: Department of Geography, University of Sheffield, Winter street, Sheffield S10 2TN, United Kingdom.

³ Current address: School of Environment, Tsinghua University, Beijing 200084, PR China.

⁴ Current address: Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK.

Supplementary material and methods

Quantitative-PCR

Q-PCR standards for each molecular target were obtained using a 10-fold serial dilution of plasmids carrying a single cloned target gene or relevant part thereof. The standards were constructed by cloning the PCR product of the environmental samples of each individual PCR assay into pCR2.1 TOPO vector by using the TOPO TA cloning kit (Invitrogen, Breda, the Netherlands) according to the manufacturer's protocol. Cloned inserts were isolated using the Qiagen Plasmid mini Kit and checked for concentration and purity on a Nanodrop ND-1000 spectrophotometer (Isogen). Presence of the gene of interest was confirmed by sequence-analysis (MWG-Biotech, Germany). The total number of plasmids with cloned target genes in the Q-PCR Standard was calculated based on its total DNA concentration (Nanodrop), assuming an average molecular mass for each nucleotide pair of 660 pg/ml (Smith et al., 2006).

Standard curve template DNA and the "no template control" (NTC) were amplified in duplicate in the same plate as the environmental samples. Five tenfold dilutions were used for each Q-PCR assay. Q-PCR amplifications were performed in 25 μ l volumes containing 12.5 μ l of iQ™ SYBR® Green Supermix (Bio-Rad, Hemel Hempstead, UK), 8.5 μ l of nuclease-free water (Ambion, Warrington, UK), 1.25 μ l of each primer (10 μ M) and 1 μ l of template DNA using a CFX96™ Real-Time System (Bio-Rad, Hemel Hempstead, UK). Standard amplification was used for all Q-PCR assays except archaeal *amoA*, starting with an initial denaturation at 95 °C for 3 min, followed by 40 cycles of 30 s at 95 °C, 0.5 to 1 min of annealing (annealing temperature and time for each primers pairs are given in Table S1), and 30 s at 72 °C. Amplification for the archaeal *amoA* gene followed the procedure as described by (Tsiknia et al., 2013). The fluorescence was measured at the end of each synthesis step (i.e. at 81 °C for archaeal *amoA* and at 72 °C for all other genes)

Threshold cycle (Ct) values and amplicon numbers were determined automatically using the Bio-Rad CFX Manager™ software. The efficiency of the Q-PCR assays was above 90%,

except for fungi and AOA (~70%). The r^2 were > 0.99 , except for *nifH* and *nosZ* genes (~0.97). Specificity of the Q-PCR was assessed via a melting curve analysis (increase of temperature from annealing temperature to 95 °C by 0.5 °C per step of 0.05 s) at the end of each Q-PCR amplification (Ririe et al., 1997). The melting curves for the bacterial and archaeal 16S rRNA, *nifH*, *amoA*, *narG*, *nirS*, and *nosZ* genes Q-PCR assays showed specificity for the amplified targeted genes (i.e. single peak). As expected, the melting curve of the Q-PCR for fungal ITS showed the amplification of products of different lengths, due to the variability in length of ITS regions among different fungal taxa (Manter and Vivanco, 2007).

Microbial respiration

Greenhouse gas fluxes from the aggregate size fractions and the bulk soil were measured from field moist bulk soil and soil aggregates (pF 3.8 -4.0; hereafter named “field moisture”) and from moistened samples (40 – 60 % of field capacity) by adding distilled water 48 hours before flux measurements started (hereafter named “elevated moisture”). Soil temperature was set to 20 °C. The soil moisture was increased because at the time of soil sampling the soil moisture content was low (pF 3.8-4.0), potentially reducing microbial activity and subsequent GHG fluxes.

Fluxes of CO₂ and NO were measured with a fully automated laboratory measuring system with 13 adapted Kilner jars serving as test chambers in a temperature-controlled incubator and connected to a CO₂ and a NO_x analyser. Twelve test chambers were used as incubation chambers. One chamber was used as a reference where no soil was incubated. The measuring system is described in detail by Schindlbacher et al. (2004) and Schaufler et al., (2010). For CO₂ and NO flux determination, air from inside the incubator was drawn through the chambers to the CO₂ and NO_x analysers with a constant flow rate of 1.0 l min⁻¹. To avoid accumulation of CO₂ and NO in the incubator, the incubator was flushed with compressed ambient air (1.0 l min⁻¹). Carbon dioxide was measured with a PP Systems WMA-2 (Amesbury, MA, USA), infrared CO₂ analyser, and NO was measured with a HORIBA APNA-360 (Kyoto, Japan)

chemoluminescence NO_x analyser. The measuring time of each chamber was 8 min according to achievement of steady state.

Determination of N₂O and CH₄ fluxes was done manually by closed chamber technique. The soil samples were put into Kilner jars and closed air-tight with a PVC lid. A glass tube, with a total volume of 685 cm³, was fitted into the PVC lid and closed air-tight with rubber septa and sealed with silicon grease. Twelve ml of the gas sample were extracted from each chamber in triplicate at intervals of 15min and injected into sealed and pre-evacuated sampling vials with a glass syringe. The analysis was done immediately by gas chromatography (AGILENT 6890N) connected to an automated system sample-injection (AGILENT TECH G1888, Network HEADSPACE-SAMPLER) at an oven temperature of 40 °C. Nitrous oxide was measured by a ⁶³Ni-electron-capture detector (ECD; detector: 350 °C) and CH₄ by a flame ionization detector (FID; detector: 250 °C). Standard gases (Inc. Linde Gas) were used as a reference and contained 0.5, 1 and 2.5 µl l⁻¹ N₂O; 1, 2 and 4µl l⁻¹ CH₄. Data were calculated as described in Kitzler et al. (2006).

Physico-chemical analysis of bulk soil and aggregates

The moisture content of each aggregate size class and the bulk soil was measured gravimetrically at 105 °C. The mass distribution over the predefined aggregate size classes was obtained by dry sieving of 100 g bulk soil from each sampling spot in triplicate (i.e. 9 replicates per site). Particle size distribution (i.e. the fractions of sand, silt and clay) for each aggregate size class and the bulk soil was determined by wet-sieving (20–2000 µm fractions) and sedimentation of the < 20 µm fraction in an X-ray sedigraph (Micromeritics Sedigraph 5000ET) after removal of organic matter with hydrogen peroxide and dispersion with sodium polyphosphate (Soil Survey Staff, 2004).

Total carbon was quantified by dry combustion (Tabatabai and Bremner, 1991) in an elemental analyser (Carlo Erba Nitrogen Analyser 500, Milano, Italy), and carbonate was measured gas-volumetrically (Soil Survey Staff, 2004). Soil organic C (SOC) was calculated as the difference of total and carbonate C. Soil and aggregate samples were extracted for N-NO₃⁻, N-

NH_4^- , and P-PO_4^- using 2 g of soil and 20 ml of KCl (1 M) shaken for 1 h. Concentration of N-NO_3^- was determined by the vanadium reduction method (Miranda et al., 2001), concentration of N-NH_4^- by the sodium salicylate-sodium nitroprusside method (Rowland, 1983), and the P-PO_4^- concentration by the ammonium molybdate–ascorbic acid method (Olsen et al., 1954).

Three different fractions of soil organic matter (SOM) were determined by simultaneous thermal analysis (STA) according to Barros et al. (2007), using 50 mg of oven dried (60 °C) samples (Netzsch STA 409 PC). The samples were heated from 25 to 600 °C at a rate of 5 °C min⁻¹ in a reaction atmosphere of synthetic air (flow rate: 50 mL min⁻¹). According to De la Rosa et al. (2008) STA allows the distinction of the amount of total SOM (decomposes between 190 and 550 °C), into thermally labile SOM (decomposes between 190 and 390 °C), thermally more stable SOM (decomposes between 390 and 450 °C), and refractory SOM (decomposes between 450 and 550 °C). In the labile fraction, SOM consists mainly of carbohydrates and proteins (De la Rosa et al., 2008), whereas in the thermally more stable SOM fraction polyphenolic and aromatic organic structures dominate (Lopes-Capel et al., 2005). Black carbon present in soil burns at higher temperatures within the refractory fraction (De la Rosa et al., 2008).

Particle size distribution in the various aggregate size classes as well as the SOM fractions (STA) were measured on one composite sample for each site (i.e. mixture of the 3 replicates/sampling spot at each site).

Table S1. Description of the primers used to target each community and the annealing temperature of each Q-PCR assays.

Target gene	Primer	Sequence 5'-3'	Annealing temp. (°C) and time (s)	References
Bacterial <i>16SrRNA</i>	519F	GCCAGCAGCCGCGGTAAT	58 (30 s)	Lane (1991); Stubner and Meuser (2000)
	907R	CCGTCAATTCCTTTGAGTTT		
Archaeal <i>16SrRNA</i>	Arch 0025F	CTGGTTGATCCTGCCAG	58 (30 s)	Vetriani et al. (1999)
	Arch 364R	ACGGGGCGCACGAGGCGCGA		
Fungal <i>ITS</i>	ITS1f	TCCGTAGGTGAACCTGCGG	50 (45 s)	Gardes and Bruns (1993); Vilgalys and Hester (1990)
	5.8s	CGCTGCGTTCTTCATCG		
<i>nifH</i>	nifHF	AAAGGYGGWATCGGYAARTCCACCAC	62.5 (60 s)	Rösch and Bothe (2005)
	nifHRb	TGSGCYTTGTCYTCRCGGATBGGCAT		
Bacteria <i>amoA</i>	amoA_F	GGHACTGGGAYTTCTGG	55.3 (30 s)	Holmes et al. (1995); Okano et al. (2004)
	amoA_R	CCTCKGSAAAGCCTTCTTC		
Archaea <i>amoA</i>	amoAF	STAATGGTCTGGCTTAGACG	55 (35 s)	Francis et al. (2005)
	amoAR	GCGGCCATCCATCTGTATGT		
<i>narG</i>	NARG F	TCGCCSATYCCGGCSATGTC	63 (30 s)	López-Gutiérrez et al. (2004)
	NARG R	GAGTTGTACCAGTCRGC SGAYTC SG		
<i>nirS</i>	NIRS4Q F	G TSAACGYSAAGGARACSGG	63 (30 s)	Braker et al. (1998)
	NIRS6Q R	GASTTCGGRTGSGTCTTSAYGAA		
<i>nosZ</i>	nosZ1840_F	CGCRACGGCAASAAGG TSMSSGT	67 (30 s)	Henry et al. (2006)
	nosZ2090_R	CAKRTGCAKSGCRTGGCAGAA		

Table S2. Overview table of the two-way ANOVA with land use and aggregate size as factors.

Significant *P* values ($P < 0.05$) are shown in bold.

	Land use		Aggregate size		Interaction	
	F values	<i>P</i> values	F values	<i>P</i> values	F values	<i>P</i> values
Bacteria	54.458	$< 2 \times 10^{-16}$	4.154	0.00161	2.754	0.00197
Archaea	9.878	2.51×10^{-5}	0.963	0.459	0.806	0.685
Fungi	9.768	2.79×10^{-5}	1.594	0.166	0.830	0.6559
<i>nifH</i>	97.755	$< 2 \times 10^{-16}$	1.635	0.155	1.535	0.112
AOB	16.231	1.04×10^{-7}	1.275	0.28353	2.473	0.00511
AOA	88.972	$< 2 \times 10^{-16}$	0.432	0.855	1.004	0.470
<i>narG</i>	184.079	$< 2 \times 10^{-16}$	2.843	0.017331	3.314	0.000305
<i>nirS</i>	246.065	$< 2 \times 10^{-16}$	0.768	0.5986	2.045	0.0216
<i>nosZ</i>	73.592	$< 2 \times 10^{-16}$	4.694	0.00062	1.889	0.03633

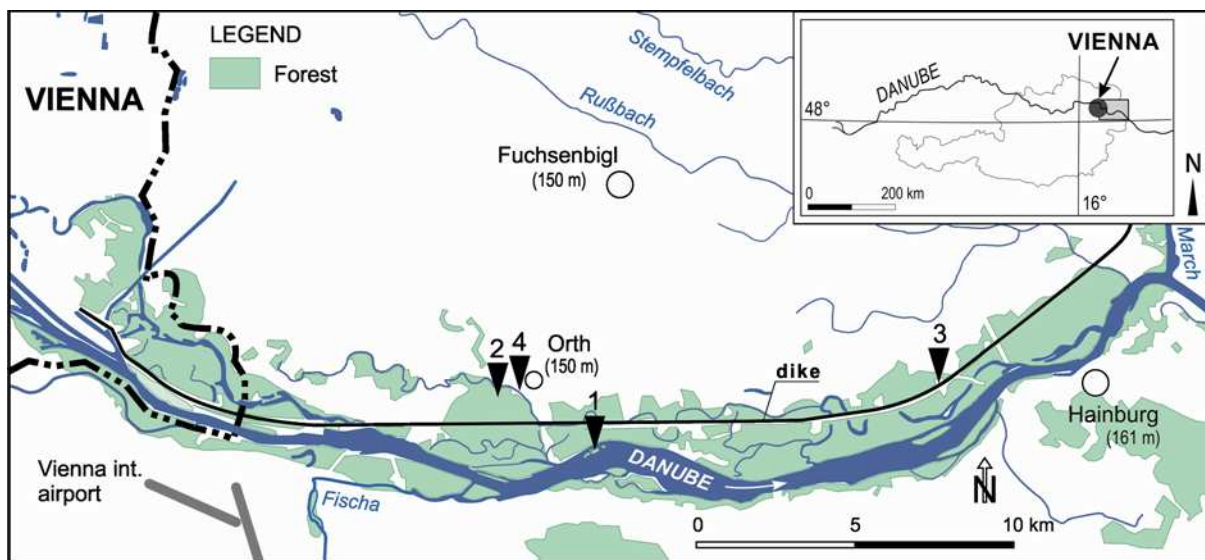


Fig. S1. Study area in the National Park “Donau-Auen” east of Vienna. The continuous black line represents a dike built from 1882 to 1905 to prevent flooding of the enclosed land. The numbers 1 to 4 indicate the 4 field sites/land uses: site 1: young forest; site 2: Forest, site 3: Grassland; site 4: Cropland.

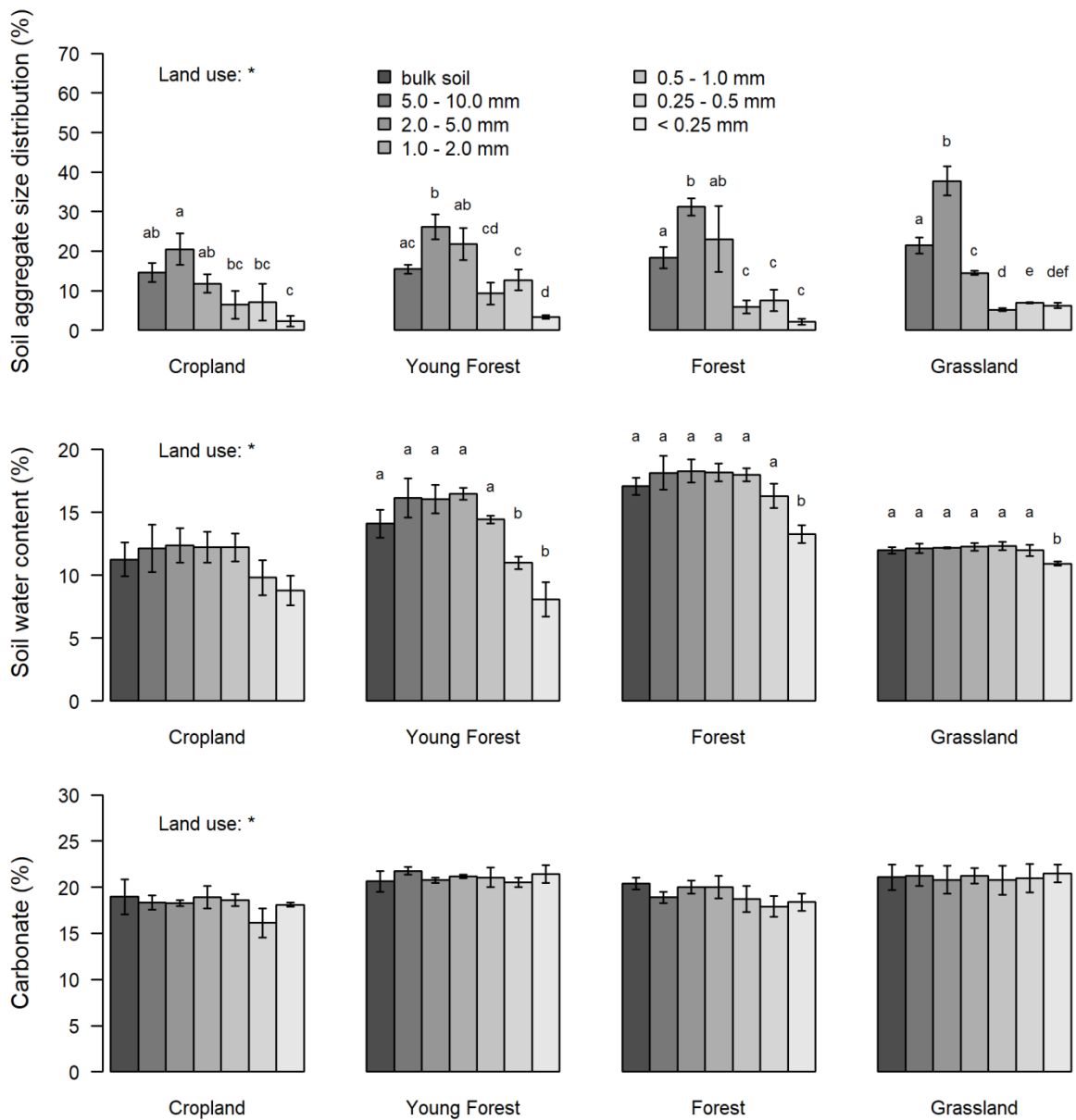


Fig. S2. Variation in soil aggregate size distribution (%), soil water content (%), and carbonate concentration (%) between bulk soil and six soil aggregates sizes classes from four land use types. Mean value \pm one standard deviation ($n = 3$) are shown. Land use: * indicates significant ($P < 0.05$) effect of land use. Small letters indicate significance ($P < 0.05$) of pairwise differences between soil aggregate size classes within a specific land use.

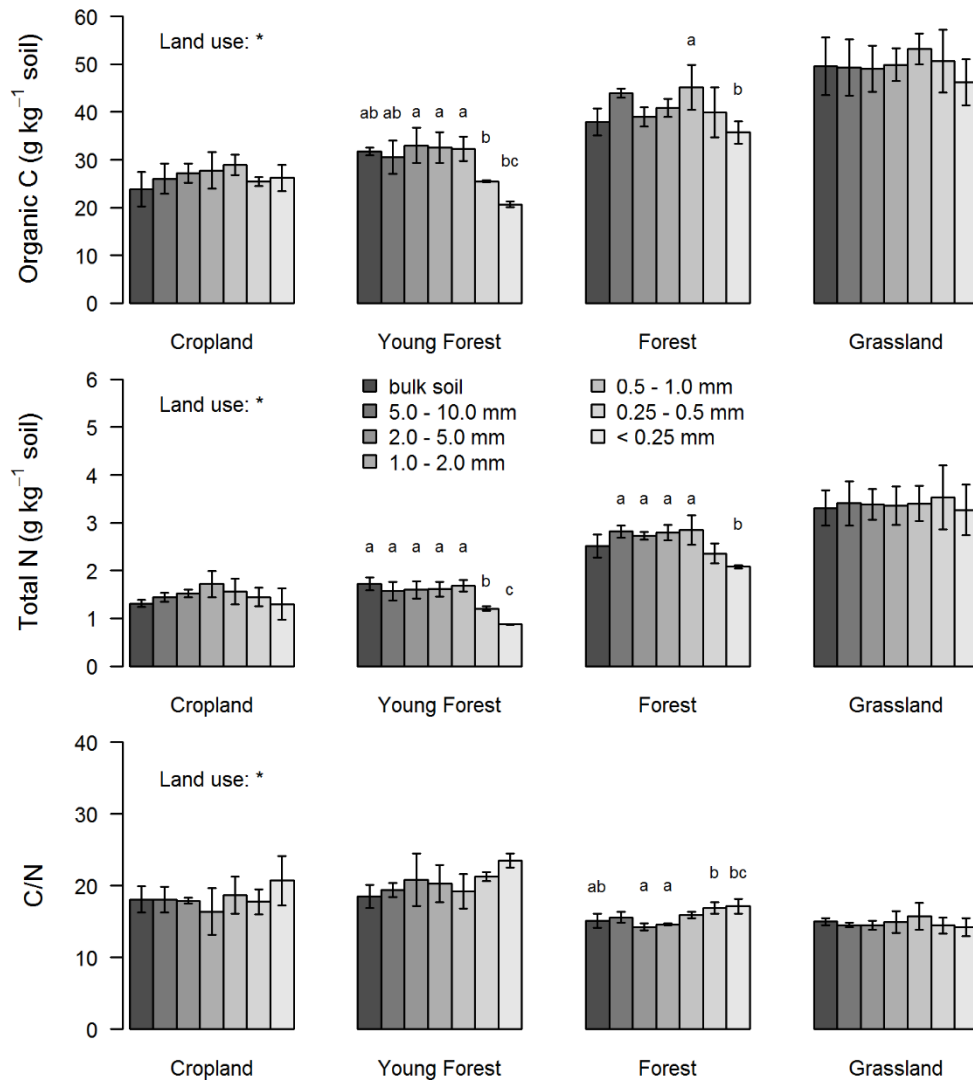


Fig. S3. Variation in organic C (g kg⁻¹ soil) and total N (g kg⁻¹ soil) concentration and C/N ratio between bulk soil and six soil aggregate sizes classes from four land use types. Mean value \pm one standard deviation ($n = 3$) are shown. Land use: * indicates significant ($P < 0.05$) effect of land use. Small letters indicate significance ($P < 0.05$) of pairwise differences between soil aggregate size classes within a specific land use.

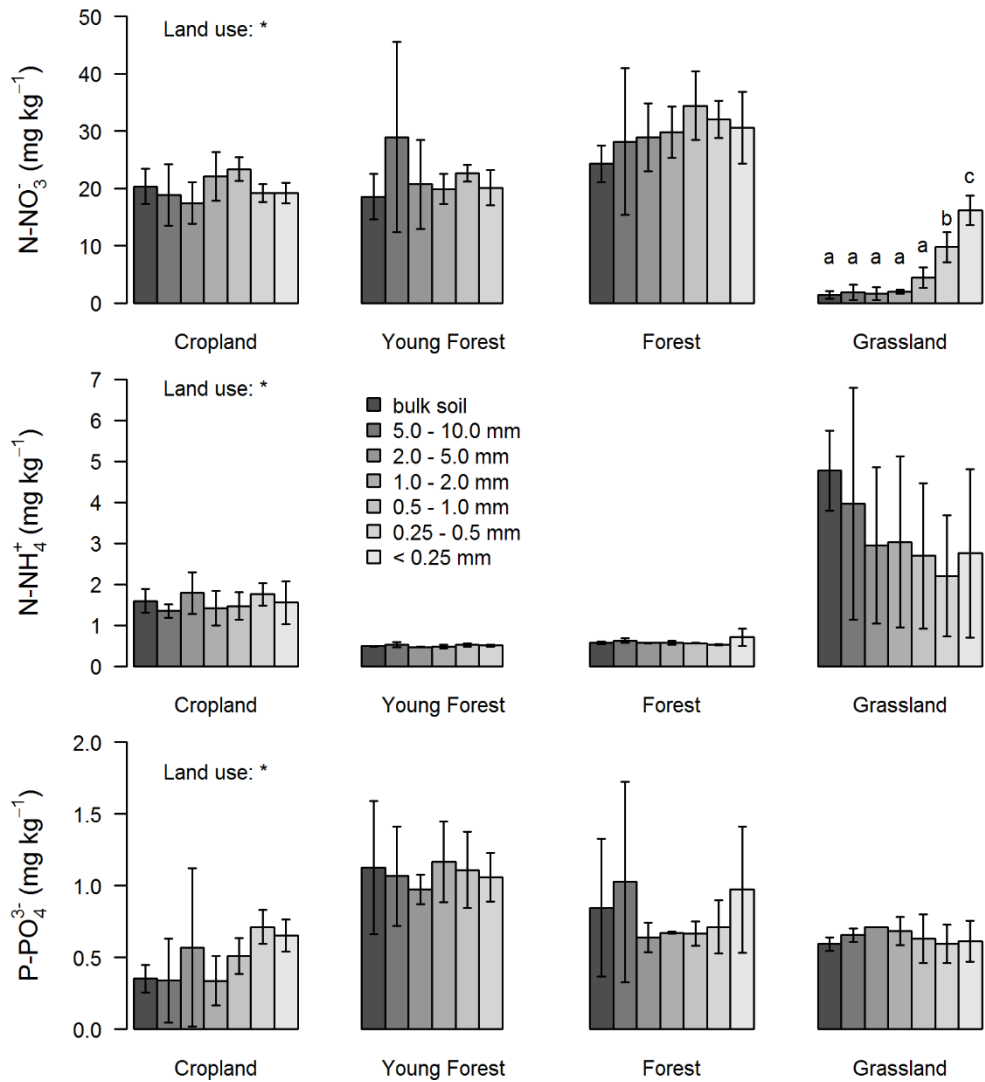


Fig. S4. Variation in N-NO_3^- (mg kg^{-1} soil), N-NH_4^+ (mg kg^{-1} soil) and P-PO_4^{3-} (mg kg^{-1} soil) concentrations between bulk soil and six soil aggregate sizes classes from four land use types. Mean value \pm one standard deviation ($n = 3$) are shown. Land use: * indicates significant ($P < 0.05$) effect of land use. Small letters indicate significance ($P < 0.05$) of pairwise differences between soil aggregate size classes within a specific land use. The N-NO_3^- , N-NH_4^+ and P-PO_4^{3-} concentrations were not measured on the < 0.25 mm aggregates from young forest site.

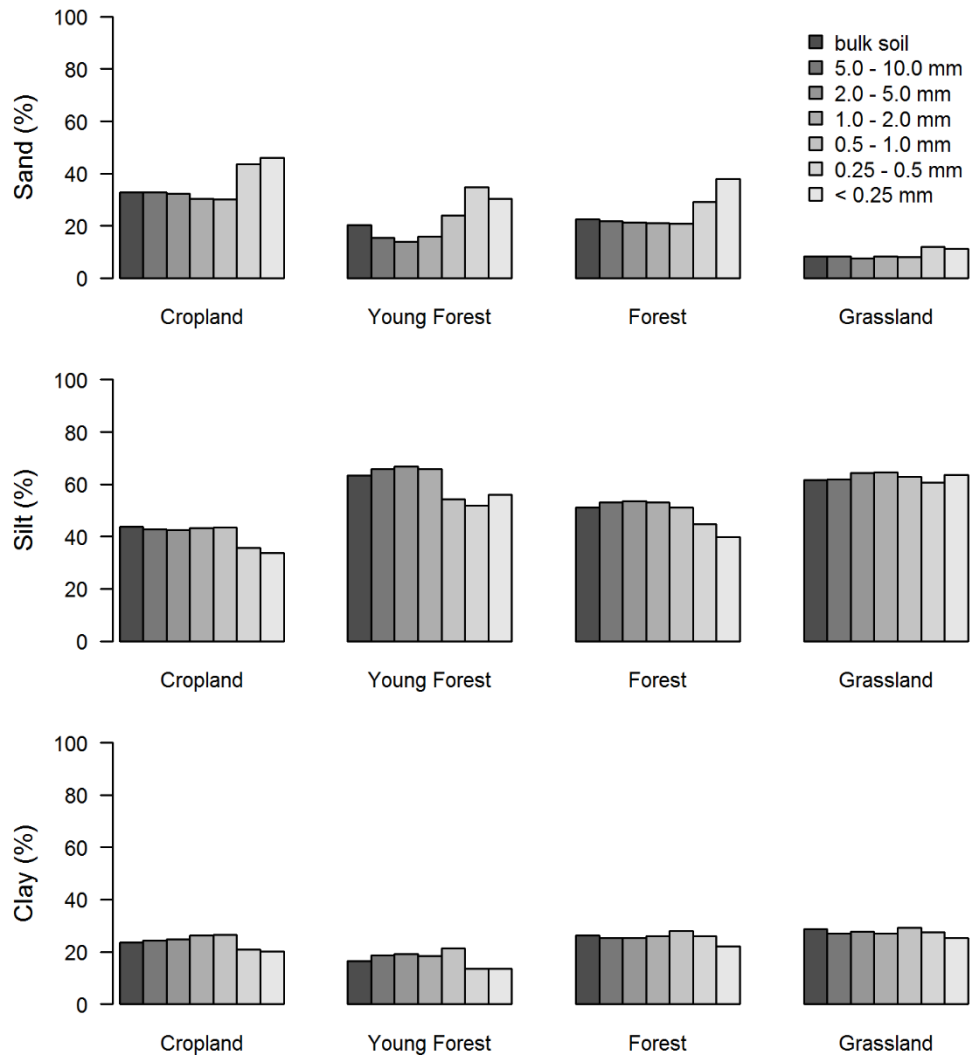


Fig. S5. Variation in sand, silt and clay contents (%) between bulk soil and six soil aggregates sizes classes from four land use types. The measurements were performed on one composite sample (mixture of 3 soil replicates).

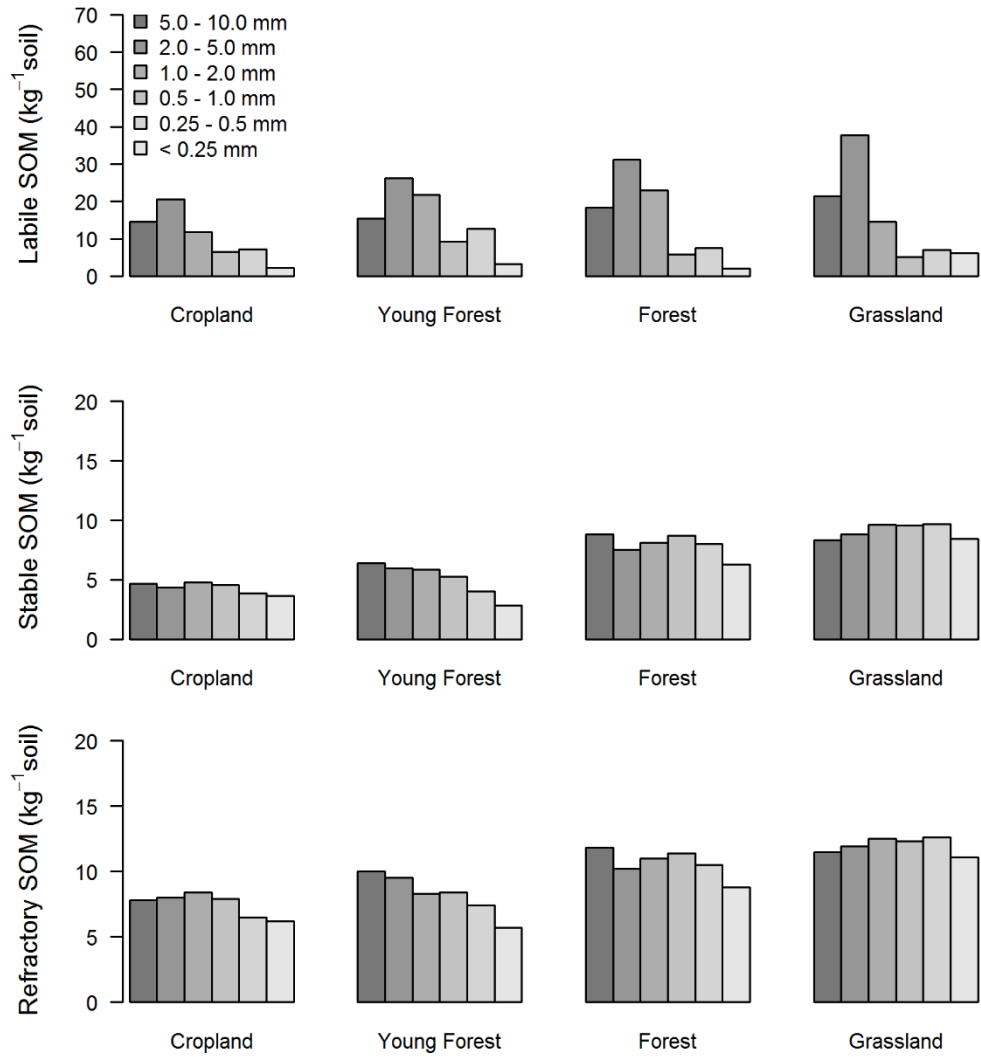


Fig. S6. Variation in labile, stable and refractory soil organic matter (SOM; g kg⁻¹ soil) between bulk soil and six soil aggregates sizes classes from four land use types. The measurements were performed on one composite sample (mixture of 3 soil replicates).

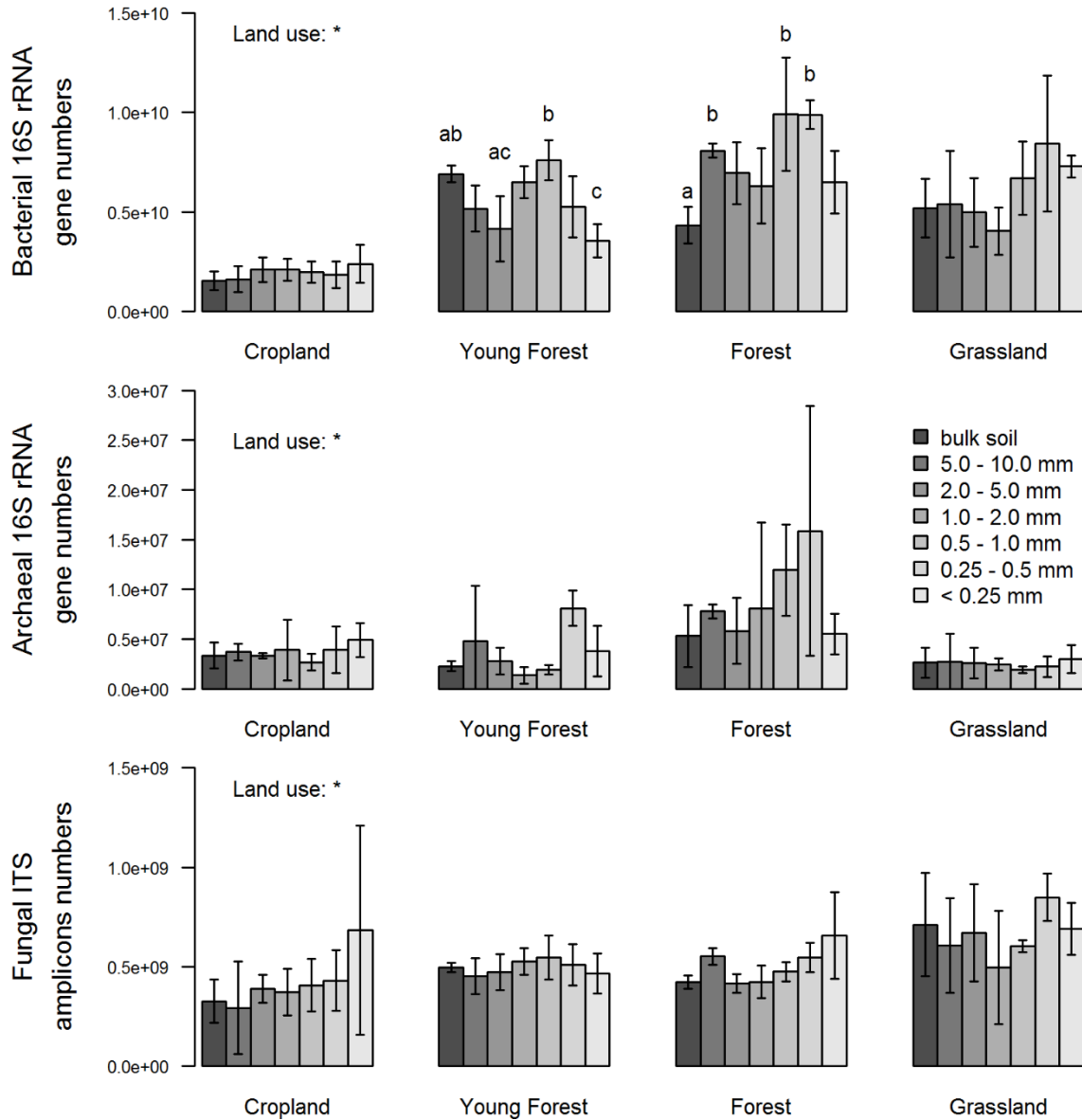


Fig. S7. Variation in gene abundance of bacteria and archaea (16S rRNA gene) and fungi (ITS amplicon) between bulk soil and 6 different soil aggregates sizes classes from 4 different land uses. The abundances of microbial communities are expressed by g^{-1} dry soil aggregates or by g^{-1} dry soil for the bulk soil. Means values \pm standard deviation ($n = 3$) are shown. Land use: * indicates significant ($P < 0.05$) effect of land use on microbial gene abundance. Different minuscule letters indicate significant ($P < 0.05$) differences between soil aggregates sizes classes for a specific land use.

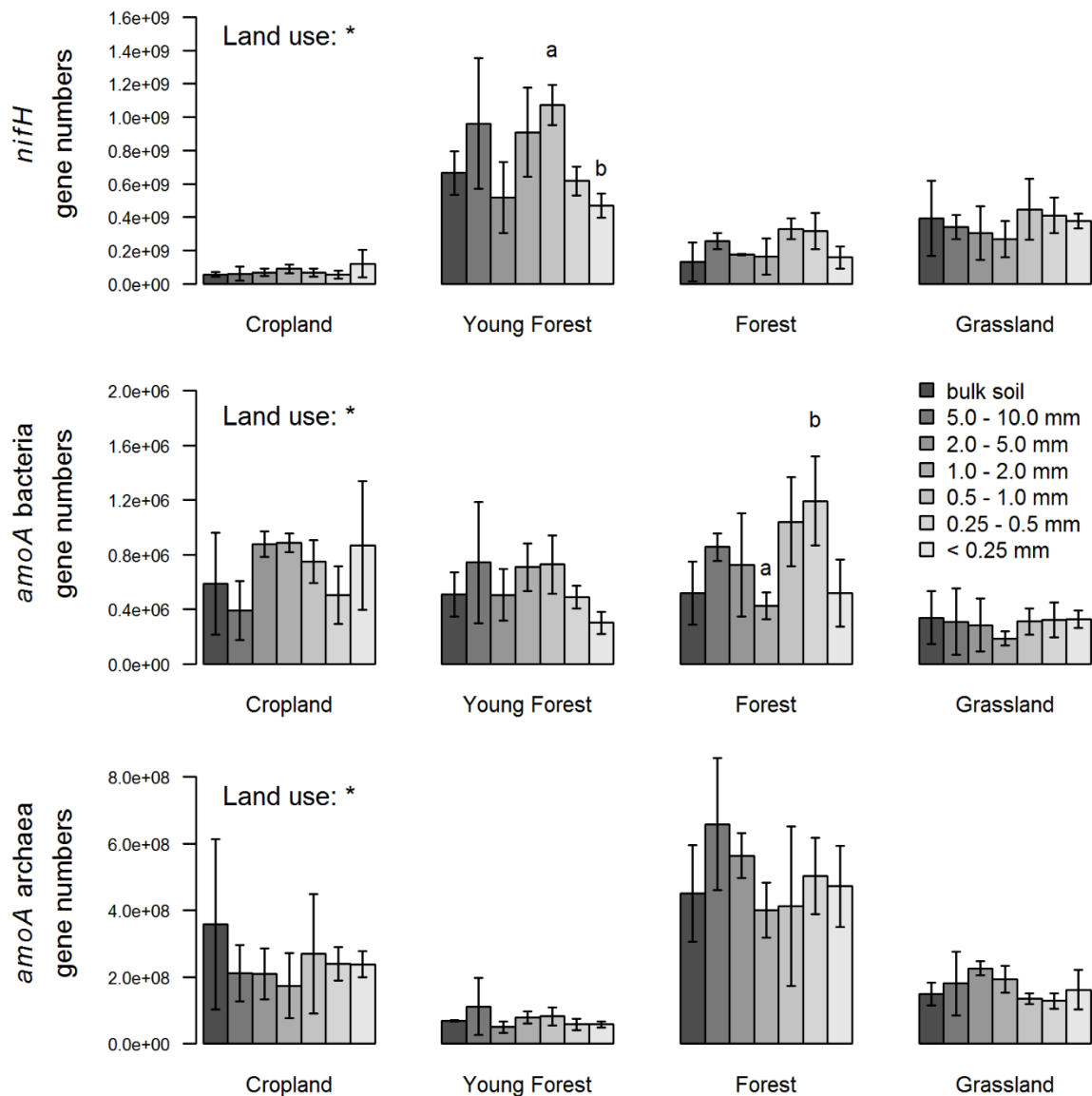


Fig. S8. Variation in gene abundance of N fixation (*nifH* gene) and ammonia oxidizing bacteria and archaea (*amoA* gene) between bulk soil and 6 different soil aggregates sizes classes from 4 different land uses. The abundances of microbial communities are expressed by g^{-1} dry soil aggregates or by g^{-1} dry soil for the bulk soil. Mean values \pm standard deviation ($n = 3$) are shown. Land use: * indicates significant ($P < 0.05$) effect of land use on microbial gene abundance. Different minuscule letters indicate significant ($P < 0.05$) differences between soil aggregate sizes classes for a specific land use.

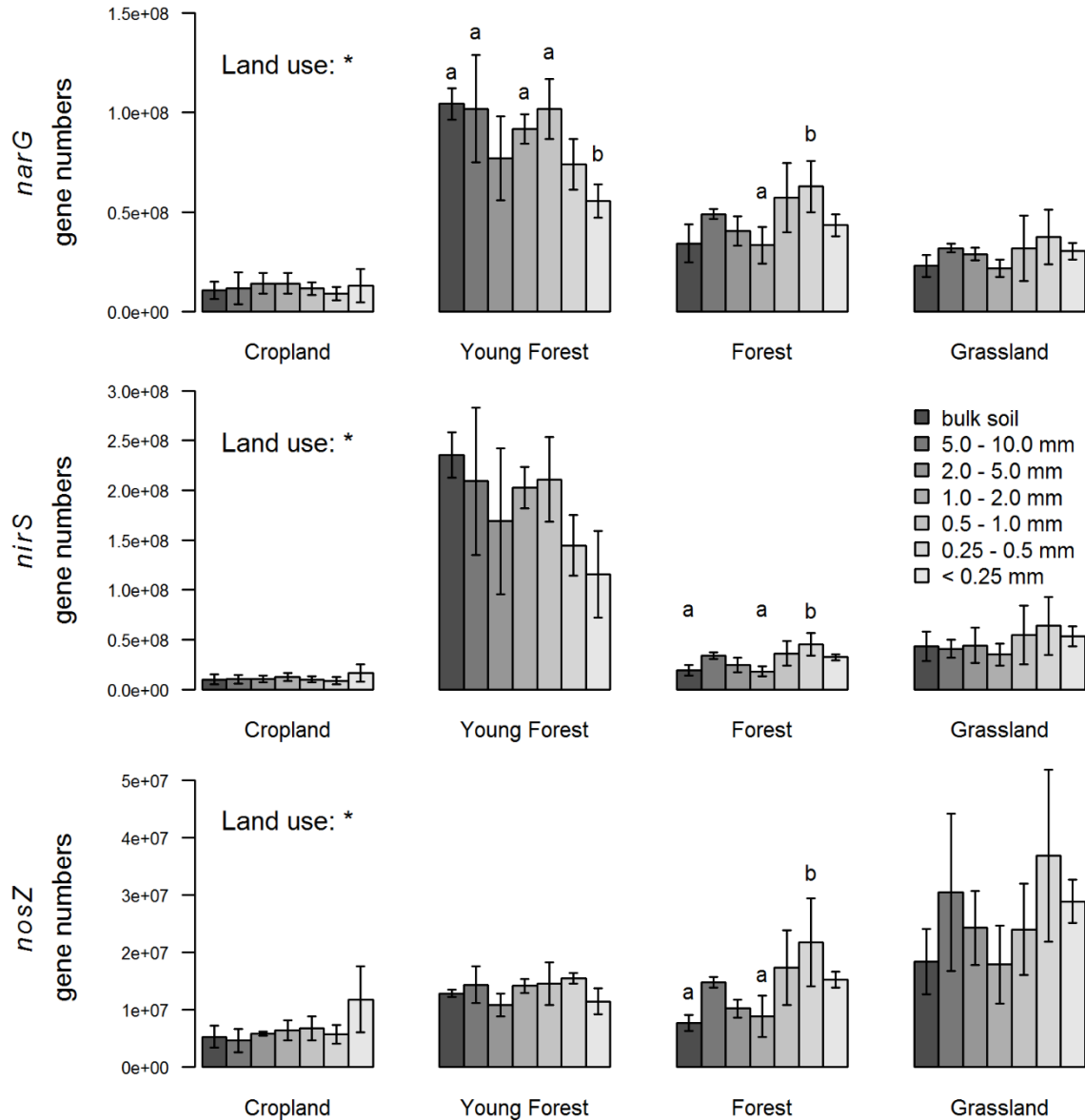


Fig. S9. Variation in gene abundance of nitrate reductase (*narG* gene), nitrite reductase (*nirK* gene) and nitrous oxide reductase (*nosZ* gene) between bulk soil and 6 different soil aggregates sizes classes from 4 different land uses. The abundances of microbial communities are expressed by g^{-1} dry soil aggregates or by g^{-1} dry soil for the bulk soil. Means values \pm standard deviation ($n = 3$; except for *nosZ* gene from cropland of the 1.0 – 2.0 mm soil aggregates, for which $n = 2$) are shown. Land use: * indicates significant ($P < 0.05$) effect of land use on microbial gene abundance. Different minuscule letters indicate significant ($P < 0.05$) differences between soil aggregates sizes classes for a specific land use.

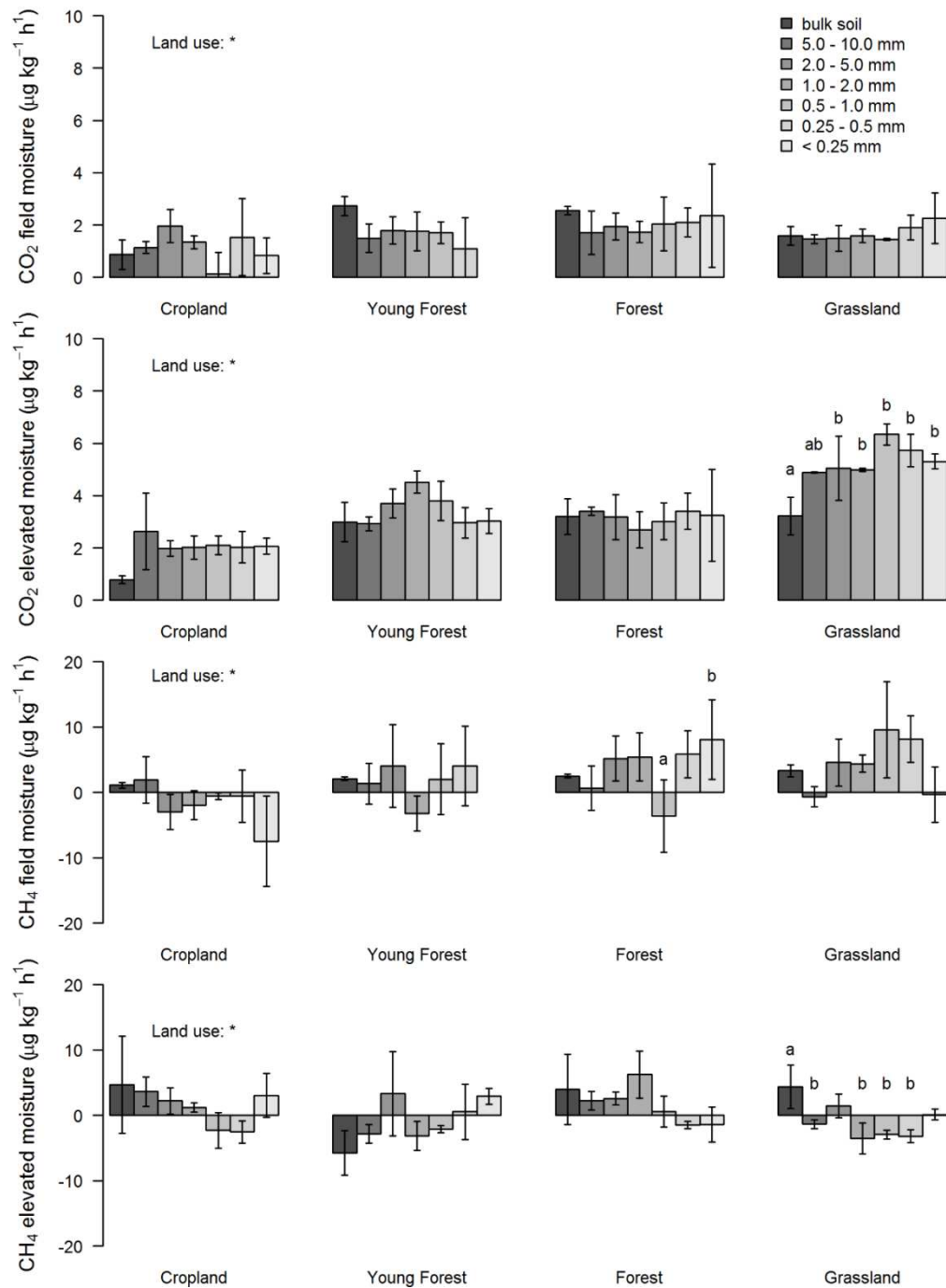


Fig. S10. Variation in CO₂ and CH₄ production ($\mu\text{g kg}^{-1} \text{h}^{-1}$) between 6 sizes fractions and bulk soil, from 4 different land uses at the field moisture or elevated moisture (40 – 60 % of field capacity). Means values \pm standard deviation ($n = 3$). Land use: * indicates significant ($P < 0.05$) effect of land use on microbial gene abundance. Different minuscule letters indicate significant ($P < 0.05$) differences between soil aggregates sizes for a specific land use. The CO₂ and CH₄ emissions were not measured for the < 0.25 mm soil fractions from young forest site at field moisture.

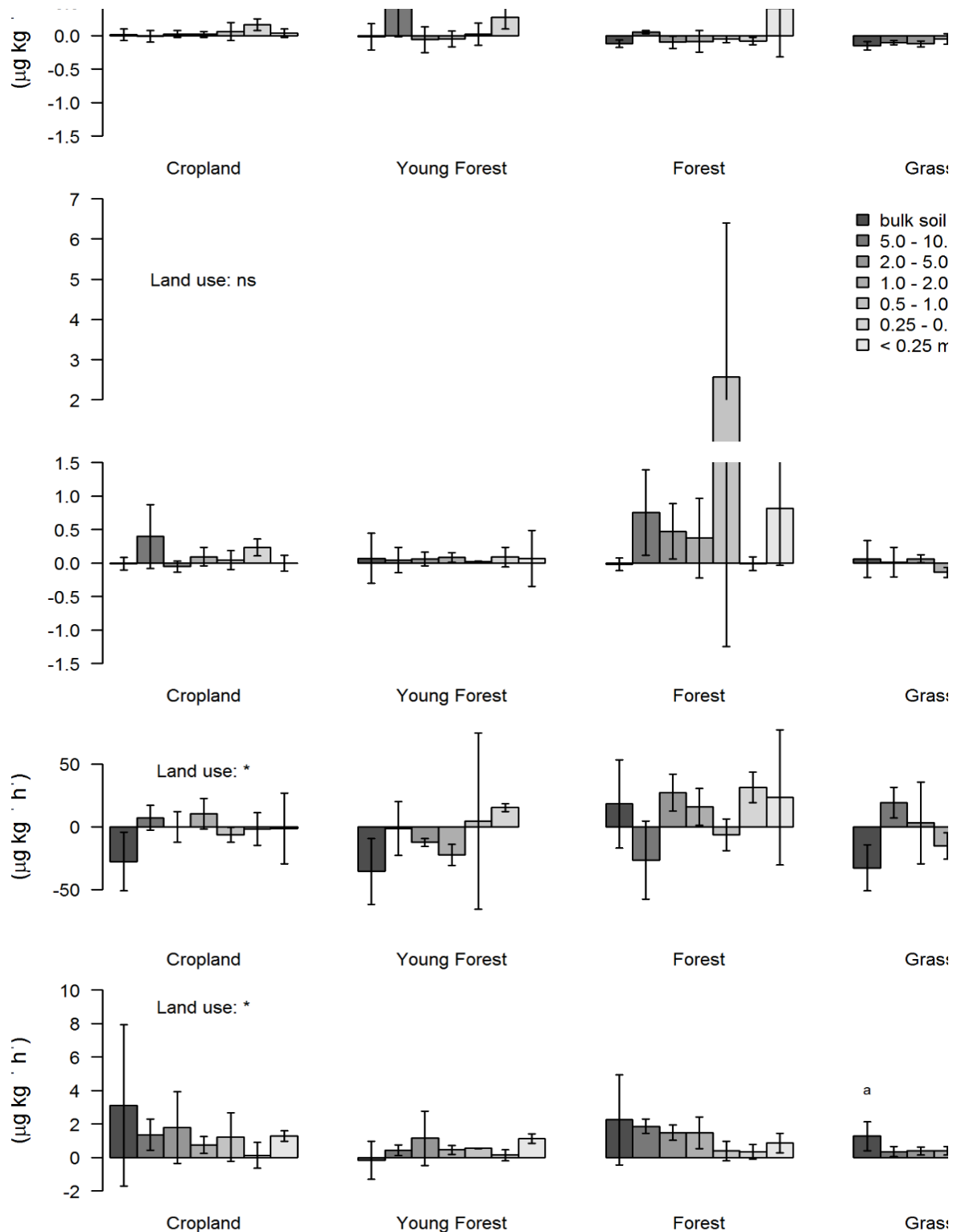


Fig. S11. Variation in NO and N₂O production ($\mu\text{g kg}^{-1} \text{ h}^{-1}$) between 6 sizes fractions and bulk soil, from 4 different land uses at the field moisture or elevated moisture (40 – 60 % of field capacity). Means values \pm standard deviation ($n = 3$). Land use: indicates significant (*: $P < 0.05$) or no (ns: non-significant $P > 0.05$) effect of land use on microbial gene abundance. Different minuscule letters indicate significant ($P < 0.05$) differences between soil aggregates sizes for a specific land use. The NO and N₂O emissions were not measured for the < 0.25 mm soil fractions from young forest site at field moisture. NB: the y-scale of N₂O is different between plots based on field moisture or elevated soil moisture.

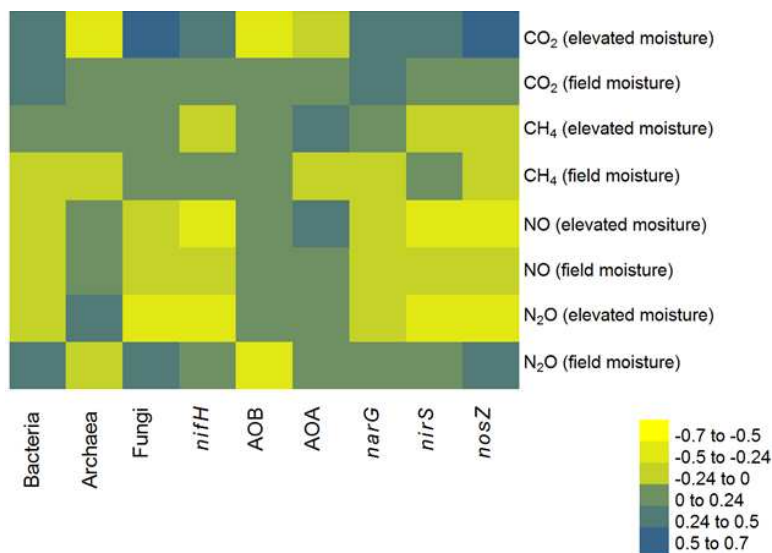


Fig. S12. Heatmaps of Spearman's rank correlation coefficients ρ between microbial genes abundance and greenhouse gas fluxes from samples across six soil aggregates sizes classes (< 0.25, 0.25 – 0.5, 0.5 – 1.0, 1.0 – 2.0, 2.0 – 5.0 and 5.0 – 10.0 mm) and four land uses. AOB: *amoA* bacteria; AOA: *amoA* archaea. The ρ values > 0.24 and < -0.24 are significant ($P < 0.05$).

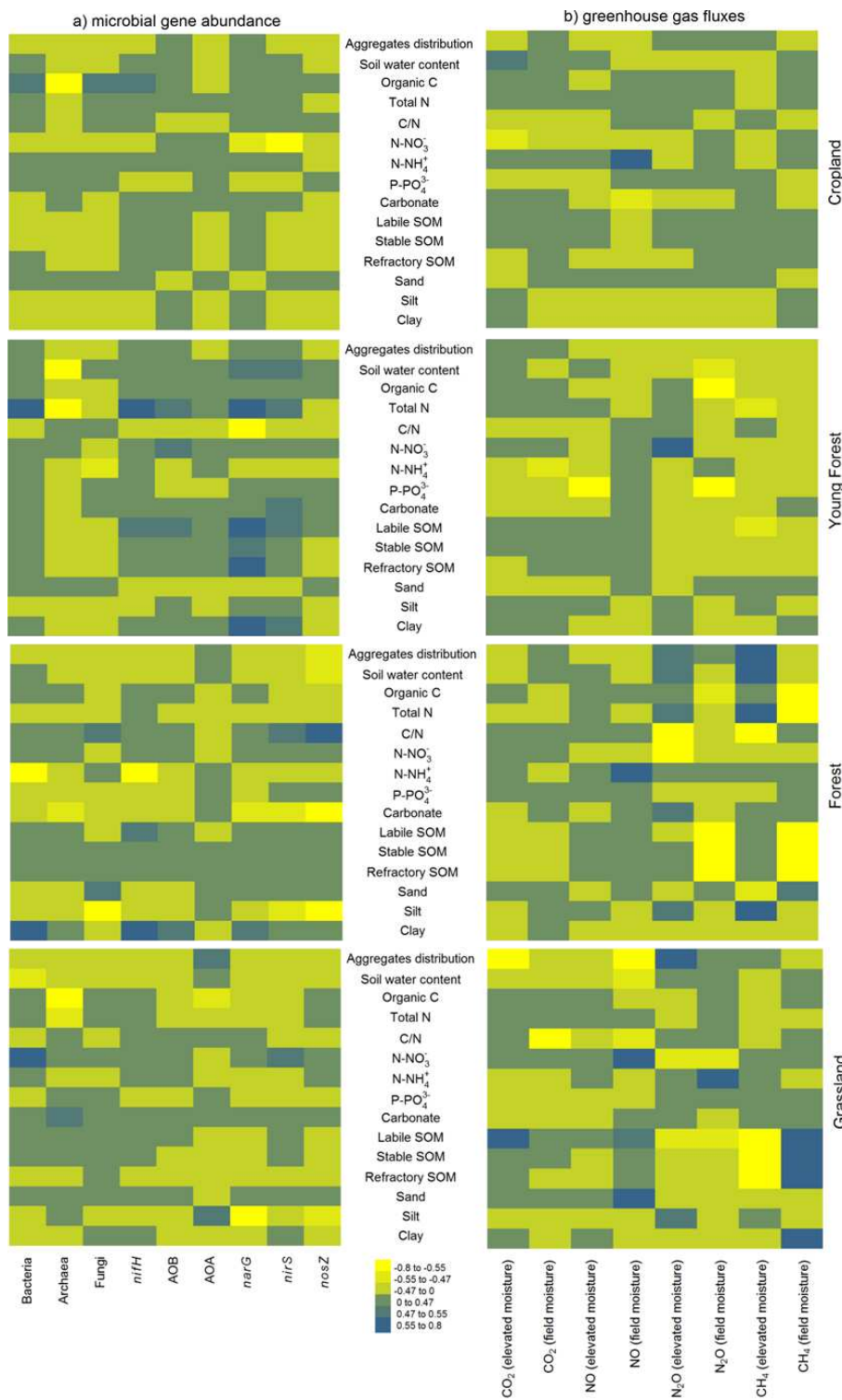


Fig. S13. Heatmaps of Spearman's rank correlation coefficients ρ between soil properties and a) microbial genes abundance or b) greenhouse gas fluxes from samples across six soil aggregates sizes classes (< 0.25 , $0.25 - 0.5$, $0.5 - 1.0$, $1.0 - 2.0$, $2.0 - 5.0$ and $5.0 - 10.0$ mm) and for four land uses separately. AOB: *amoA* bacteria; AOA: *amoA* archaea. The ρ values > 0.47 and < -0.47 are significant ($P < 0.05$).

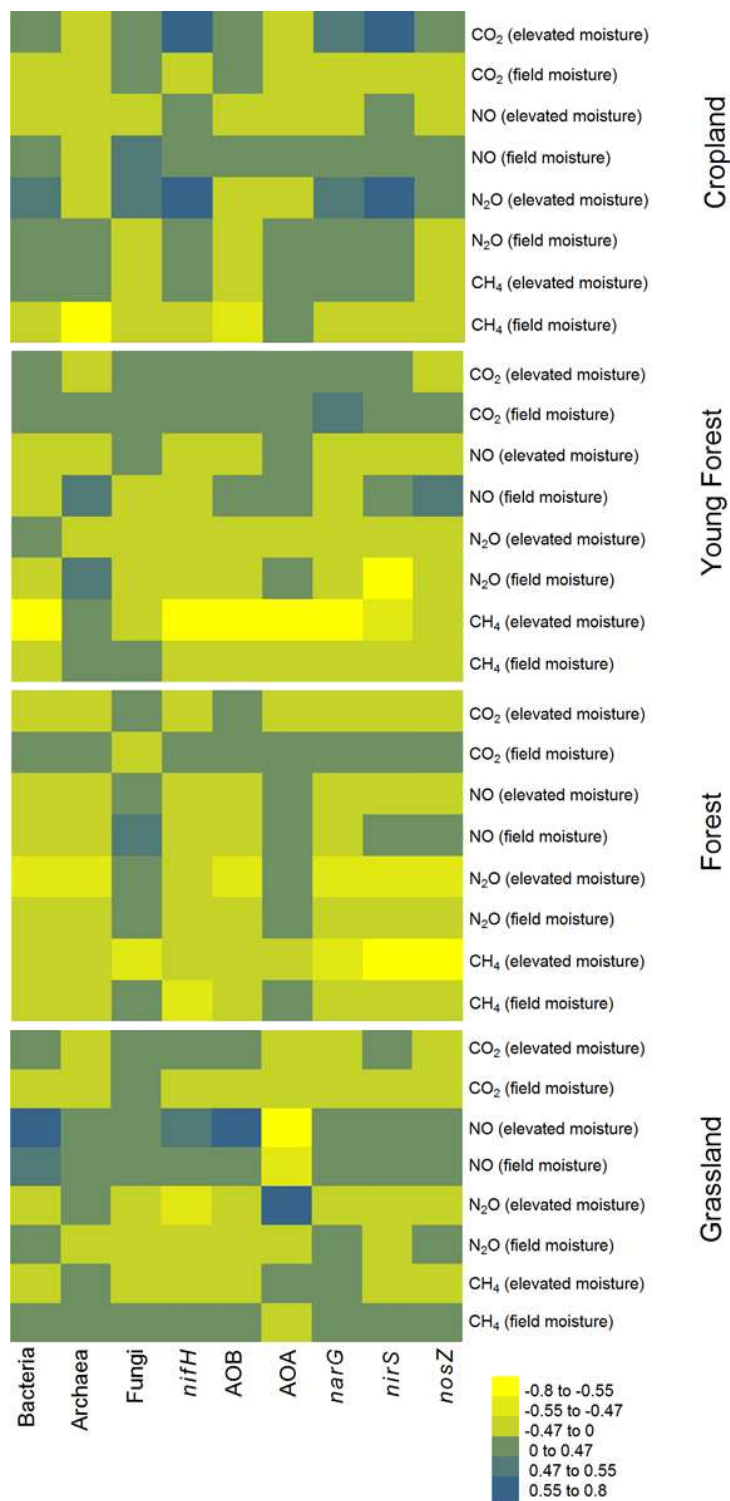


Fig. S14. Heatmaps of Spearman's rank correlation coefficients ρ between microbial genes abundance and greenhouse gas fluxes from samples across six soil aggregates sizes classes (< 0.25, 0.25 – 0.5, 0.5 – 1.0, 1.0 – 2.0, 2.0 – 5.0 and 5.0 – 10.0 mm) and for four land uses separately. AOB: *amoA* bacteria; AOA: *amoA* archaea. The ρ values > 0.47 and < -0.47 are significant ($P < 0.05$).

References:

- Barros, N., Salgado, J., Feijóo, S., 2007. Calorimetry and soil. *Thermochim. Acta*, XIVth ISBC Proceedings Special Issue Fourteenth conference of the International Society for Biological Calorimetry 458, 11–17.
- Braker, G., Fesefeldt, A., Witzel, K.-P., 1998. Development of PCR primer systems for amplification of nitrite reductase genes (*nirK* and *nirS*) To detect denitrifying bacteria in environmental samples. *Appl. Environ. Microbiol.* 64, 3769–3775.
- De la Rosa, J.M., Knicker, H., López-Capel, E., Manning, D.A.C., González-Perez, J.A., González-Vila, F.J., 2008. Direct detection of black carbon in soils by Py-GC/MS, carbon-13 NMR spectroscopy and thermogravimetric techniques. *Soil Sci. Soc. Am. J.* 72, 258.
- Francis, C.A., Roberts, K.J., Beman, J.M., Santoro, A.E., Oakley, B.B., 2005. Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. *Proc. Natl. Acad. Sci. U. S. A.* 102, 14683–14688.
- Gardes, M., Bruns, T.D., 1993. ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. *Mol. Ecol.* 2, 113–118.
- Henry, S., Bru, D., Stres, B., Hallet, S., Philippot, L., 2006. Quantitative detection of the *nosZ* Gene, encoding nitrous oxide reductase, and comparison of the abundances of 16S rRNA, *narG*, *nirK*, and *nosZ* genes in soils. *Appl. Environ. Microbiol.* 72, 5181–5189.
- Holmes, A.J., Costello, A., Lidstrom, M.E., Murrell, J.C., 1995. Evidence that participate methane monooxygenase and ammonia monooxygenase may be evolutionarily related. *FEMS Microbiol. Lett.* 132, 203–208.
- Kitzler, B., Zechmeister-Boltenstern, S., Holtermann, C., Skiba, U.M., Butterbach-Bahl, K., 2006. Controls over N₂O, NO_x and CO₂ fluxes in a calcareous mountain forest soil. *Biogeosciences* 3, 383–395.
- Lane, D.J., 1991. *Nucleic acid techniques in bacterial systematics*. John Wiley & Sons.

- Lopes-Capel, E., Sohi, S., Gaunt, J.L., Manning, D.A.C., 2005. Use of thermo gravimetry-differential scanning calorimetry to characterize soil organic matter fractions. *Soil Sci. Soc. Am. J.* 69, 136–140.
- López-Gutiérrez, J.C., Henry, S., Hallet, S., Martin-Laurent, F., Catroux, G., Philippot, L., 2004. Quantification of a novel group of nitrate-reducing bacteria in the environment by real-time PCR. *J. Microbiol. Methods* 57, 399–407.
- Manter, D.K., Vivanco, J.M., 2007. Use of the ITS primers, ITS1F and ITS4, to characterize fungal abundance and diversity in mixed-template samples by qPCR and length heterogeneity analysis. *J. Microbiol. Methods* 71, 7–14.
- Miranda, K.M., Espey, M.G., Wink, D.A., 2001. A Rapid, simple spectrophotometric method for simultaneous detection of nitrate and nitrite. *Nitric Oxide Biol. Chemistry* 5, 62–71.
- Okano, Y., Hristova, K., Leutenegger, C., Jackson, L.E., Denison, R.F., Gebreyesus, B., Lebauer, D., Scow, K.M., 2004. Application of real-time PCR to study effects of ammonium on population size of ammonia-oxidizing bacteria in soil. *Appl. Environ. Microbiol.* 70, 1008–1016.
- Olsen, S., Cole, C.V., Dean, L.A., Watanabe, F.S., 1954. Estimation of available phosphorus in soils by extraction with sodium bicarbonate. *USDA Circ. No 939 Wash. DC US Gov. Print. Off.*
- Ririe, K.M., Rasmussen, R.P., Wittwer, C.T., 1997. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. *Anal. Biochem.* 245, 154–160.
- Rösch, C., Bothe, H., 2005. Improved assessment of denitrifying, N₂-fixing, and total-community bacteria by terminal restriction fragment length polymorphism analysis using multiple restriction enzymes. *Appl. Environ. Microbiol.* 71, 2026–2035.
- Rowland, A.P., 1983. An automated method for the determination of ammonium-n in ecological materials. *Commun. Soil Sci. Plant Anal.* 14, 49–63. Schaufler, G., Kitzler, B., Schindlbacher, A., Skiba, U., Sutton, M.A., Zechmeister-Boltenstern, S., 2010. Greenhouse gas emissions from European soils under different land use: effects of soil moisture and temperature. *Eur. J. Soil Sci.* 61, 683–696.

- Schindlbacher, A., Zechmeister-Boltenstern, S., Butterbach-Bahl, K., 2004. Effects of soil moisture and temperature on NO, NO₂, and N₂O emissions from European forest soils. *J. Geophys. Res. Atmospheres* 109, 1–12.
- Smith, C.J., Nedwell, D.B., Dong, L.F., Osborn, A.M., 2006. Evaluation of quantitative polymerase chain reaction-based approaches for determining gene copy and gene transcript numbers in environmental samples. *Environ. Microbiol.* 8, 804–815.
- Soil Survey Staff, 2004. Keys to soil taxonomy, 10th edn, USDA – NRCS. ed. Washington, DC.
- Stubner, S., Meuser, K., 2000. Detection of *Desulfotomaculum* in an Italian rice paddy soil by 16S ribosomal nucleic acid analyses. *FEMS Microbiol. Ecol.* 34, 73–80.
- Tabatabai, M.A., Bremner, J.M., 1991. Automated instruments for determination of total carbon, nitrogen, and sulfur in soils by combustion techniques, in: Smith, K.A. (Ed.), *Soil Analysis*. New York, p. pp 261-286.
- Tsiknia, M., Tzanakakis, V.A., Paranychianakis, N.V., 2013. Insights on the role of vegetation on nitrogen cycling in effluent irrigated lands. *Appl. Soil Ecol.* 64, 104–111.
- Vetriani, C., Jannasch, H.W., MacGregor, B.J., Stahl, D.A., Reysenbach, A.-L., 1999. Population structure and phylogenetic characterization of marine benthic archaea in deep-sea sediments. *Appl. Environ. Microbiol.* 65, 4375–4384.
- Vilgalys, R., Hester, M., 1990. Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several *Cryptococcus* species. *J. Bacteriol.* 172, 4238–4246.