

Appendixes

Appendix A. Description of Soil Cartographic Units (SCUs) in the Study Area according to IGAC (2015)

- HL205 Consociation.

This unit occupies 69.6% of the hydrographic unit (HU) and is found on hill landforms, including summits and slopes. Soils in this SCU formed from glacial detritus over igneous rocks (monzogranite). Taxonomically, the unit consists of approximately 80% *Entic Humudepts* with inclusions of 20% *Acrudoxic Melanudands*. These soils are very shallow and limited by abundant rock fragments throughout the profile. They are well-drained due to their very high total porosity. Their natural fertility is low, characterized by strongly to very strongly acidic pH, extremely low available phosphorus, and very high aluminum saturation throughout the profile, often at toxic levels for most crops. These soils also have high organic matter content, high cation exchange capacity (CEC), and low bulk density in the surface horizon. The unit spans three slope classes: 27% of the area falls within slopes of 7–12%, 48% within 12–25%, and 25% within 25–50%, with all areas showing slight erosion.

- HL259 Consociation.

Covering 12.9% of the HU, this unit is also located on hills, with terrain forms including summits and slopes. The parent material consists of glacial detritus over undifferentiated sedimentary rocks (siltstones and claystones). It comprises approximately 75% *Typic Humudepts*, 15% *Pachic Humudepts*, and 10% *Fluventic Humudepts*. Soils are moderately deep, well-drained, and rich in organic matter. Like HL205, they exhibit low natural fertility due to strongly acidic pH, high aluminum saturation, and low nutrient levels. Surface stoniness of varying sizes is present across more than 15% of the SCU area. No evidence of erosion is observed. Approximately 72% of this SCU lies on slopes of 7–12%, and 28% on slopes of 12–25%.

- HT111 Consociation.

This unit covers 9.8% of the study area and is located on fluvio-glacial terraces with flat terrain, where slopes range between 7 and 12%. It is composed of 75% *Typic Dystrudepts* and 25% *Typic*

Humudepts. Soils formed from fluvioglacial deposits and are deep, well-drained, and have loam textures. They are characterized by high aluminum saturation, medium natural fertility, and no evidence of erosion.

- HV85 Complex.

This unit occupies 4.7% of the HU and is located in the floodplains along the main river channel, in a small valley landscape. Parent materials consist of colluvial-alluvial deposits. These soils occur in areas prone to frequent flooding, with slopes ranging from 1–7%. The unit comprises 40% *Aquic Humudepts*, 35% *Entic Humudepts*, and 25% *Fluventic Humudepts*. Soils are shallow to moderately deep, poorly drained, and contain abundant rock fragments throughout the profile. Their physical and chemical properties are similar to those in other SCUs in the area, particularly in terms of low natural fertility.

- EQ492 Consociation.

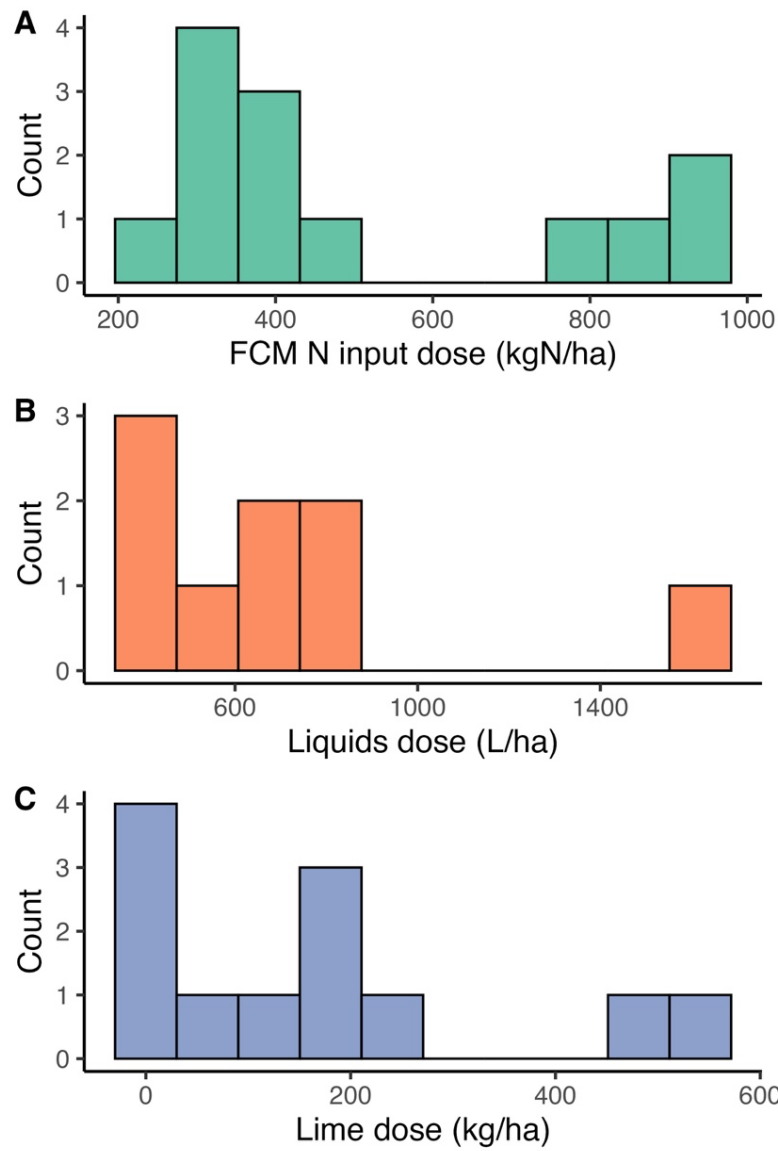
This unit covers 2.9% of the HU and is located in summit terrain with slope landforms. The parent material is biotitic monzogranite composed of potassium feldspar, plagioclase, quartz, biotite, and hornblende (Rodríguez et al., 2020). The consociation consists of approximately 75% *Andic Humicryepts*, with inclusions of 5% *Typic Humicryepts*, 5% *Lithic Melanocryands*, and 5% rock outcrops. Soils are shallow, stony, well-drained, with medium textures, high organic matter content, very high aluminum saturation, low natural fertility, and no signs of erosion. Slopes in this SCU range from 12 to 25%.

Appendix B. Survey questionnaire

Section	Questions
1. General Knowledge	How long have you lived in Santurbán-Berlín Páramo? How long have you been cultivating green onions? Have you only cultivated them in Santurbán-Berlín Páramo? How long have you been cultivating on this farm? How did you learn everything you know about cultivation? How do you evaluate the quality of your onions? Have you faced difficulties marketing onions or acquiring inputs? What are the main problems with green onion cultivation on this farm? (e.g., climate, frost, water availability/excess, pests, weeds, diseases, soil quality, input costs, etc.) Have you conducted soil tests? If so, how many tests have you conducted, and how long ago? What were the results? How much area is planted with onions on this farm?
2. Soil Fertility Perceptions and Indicators	What are the characteristics of your soil? (e.g., color, texture, etc.) How can you tell if soil is good or bad for planting? (e.g., by color, texture, stoniness, or other qualities?) Are there areas or onion plants on your farm that require more manure or chemical fertilizers or more water? If so, what might be the cause? Can you identify good and bad soil areas on your farm? Which are they? Have you noticed significant changes in soil quality or crop productivity over your time working this farm? Have you made any changes to your farming activities to improve productivity? If so, what changes? Do you think soil becomes depleted after being worked for a long time? If so, which activity contributes most to this? How do you think soil quality could be maintained?
3. Soil Preparation and Planting	How often do you prepare the soil before replanting onions? How do you prepare the soil before planting? Why do you do it this way? Do you apply lime into the soil? Why? When should lime be applied? (e.g., before or after fertilization, before or after irrigation/rainfall, once or multiple times during the crop cycle) How much lime do you apply? How do you decide how much lime to apply? How do you manage seeds before planting?
4. Fertilization and Nutrient Management	Do you use fresh chicken manure, chemical fertilizers, or other products to fertilize onions? How many of these inputs do you use for the entire farm? Why is it important (or not) to use these products? Do you treat chicken manure in any way before applying it to the crop? Do you use the same amount of those inputs in winter and summer? Are there animals, bacteria, or fungi in the soil where onions are grown? If so, which ones? Are all harmful, or can some improve soil or crop quality? Why? Have you planted or considered planting other crops on your farm besides onions? What challenges would you face in cultivating products other than onions?

Section	Questions
5. Water Management	How do you decide how much water to apply to the crop during irrigation? Can a harvest be ruined by over-irrigation? Has this happened on your farm? Does your land experience waterlogging, or does it have good drainage? If waterlogging occurs, what do you do to prevent it? Where does the water for irrigation come from? Do you use a shared tank to store irrigation water? How often do you irrigate? How many hours per day do you run the sprinklers for irrigation? What times of day do you irrigate? How many hours of irrigation are applied to each plot?
6. Pest and Diseases Management	How can you tell if an onion plant is healthy or failing? (e.g., size, height, thickness, color, root size) If an onion plant is unhealthy, what can be done to prevent the whole harvest from failing? What pests attack onion crops? How are they controlled? How often do you apply liquids for pest control? What quantities? What products? Are there new pests that weren't seen before but are now common in onions? If so, which ones and why do they appear? How do you control weeds? What do you do with the weeds you pull?
7. Acquisition of External Inputs	Do you think you spend a lot on chicken manure, lime, chemical fertilizers, liquids, or other inputs? What is the most expensive? How often do you buy chicken manure? From whom or where do you buy it? Is the quality of chicken manure consistent, or does it sometimes arrive better or worse? Where do you buy other inputs (liquids, lime, chemical fertilizers, etc.)?
8. Waste Management and Soil Contamination	Do chicken manure, chemical fertilizers, or pest control liquids contain substances that pollute the soil or harm the crop? What do you do with harvest residues? Do you think they could be used on your farm? How?

Appendix C. Histogram of input use frequency across farmers



Appendix D. Statistical analyses results for bulk density (BD) and soil water content (SWC) data across 0-5 cm and 10-15 cm depths for the initial soil monitoring campaign.

Table D1. Results of Paired T Test

Variable	t	df	p-value	Significant differences among means?
BD	-0.6018	99	0.725	No
SWC	-2.1644	91	0.9835	No

Figure D1. Boxplots for bulk density (BD) for different soil depths in Berlín Páramo

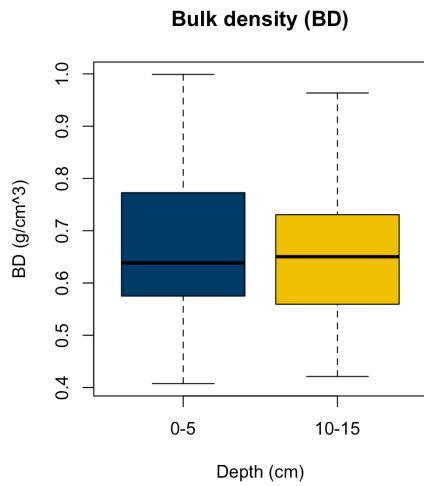
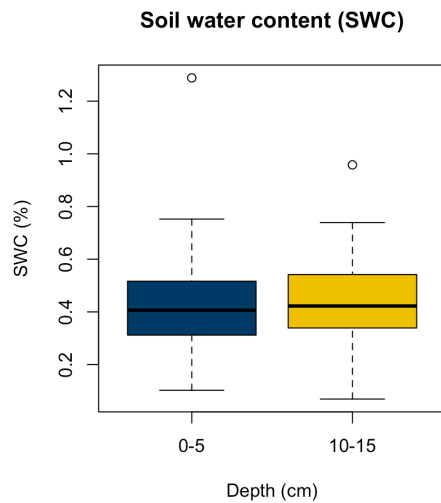


Figure D2. Boxplots for soil water content (SWC) for different soil depths in Berlín Páramo



Appendix E. Statistical analyses results for edaphic properties for the initial soil monitoring campaign.

Table E1. Results of Shapiro Wilk Normality Test

Variable	p-value	Fulfills normal distribution assumption?
TN	0.0002	No
N-NH ₄ ⁺	0.0003	No
N-NO ₃ ⁻	0.1853	Yes
pH	0.0419	No
SOC	0.0023	No
Ca ⁺²	0.0010	No
Mg ⁺²	0.0149	No
K ⁺	0.0960	No
Al ⁺³	1.354e-08	No
Na ⁺	6.947e-06	No
ECEC	0.0014	No
P	7.617e-06	No
S	0.0375	No
SWC	0.0004	No
BD	0.0570	Yes
EC	0.2968	Yes

Table E2. Results of Bartlett Test for Variables with a normal distribution

Variable	p-value	Fulfills homogeneity of variances assumption?
N-NO ₃ ⁻	0.0103	No
BD	0.5678	Yes
EC	0.3739	Yes

Appendix F. Results of statistical comparison tests for edaphic variables between cultivated and non-cultivated soils during the initial soil monitoring campaign

Variable	Test Applied	Test Statistic	p-value
BD	Welch's t-test	1.28	0.232
EC	Welch's t-test	5.66	<0.001
N-NO ₃ ⁻	Welch's t-test	9.53	<0.001
TN	Mann-Whitney U test	37.00	<0.001
N-NH ₄ ⁺	Mann-Whitney U test	190.00	0.576
SOC	Mann-Whitney U test	325.00	<0.001
pH	Mann-Whitney U test	18.00	<0.001
Ca ⁺²	Mann-Whitney U test	336.00	<0.001
Mg ⁺²	Mann-Whitney U test	335.00	<0.001
K ⁺	Mann-Whitney U test	336.00	<0.001
Al ⁺³	Mann-Whitney U test	0.00	<0.001
Na ⁺	Mann-Whitney U test	276.00	0.003
ECEC	Mann-Whitney U test	335.00	<0.001
P	Mann-Whitney U test	333.00	<0.001
S	Mann-Whitney U test	336.00	<0.001
SWC	Mann-Whitney U test	123.00	0.244

Appendix G. Mineralogical composition of soil sampling plots (XRD analysis)

Sampling plot	Depth (cm)	Minerals in bulk	Clay minerals
Native Grassland	0 – 20	Albite, Microcline, Gibbsite, Quartz, Bloedite, Calcite	Muscovite, Kaolinite, Vermiculite
	20 – 40	Albite, Microcline, Gibbsite, Quartz, Calcite	Muscovite, Vermiculite, Kaolinite
	40 – 60	Albite, Microcline, Gibbsite, Quartz, Calcite	Muscovite, Kaolinite, Vermiculite
3-Year Cultivated	0 – 20	Albite, Calcite, Gibbsite, Microcline, Quartz	Muscovite, Kaolinite, Vermiculite
	20 – 40	Albite, Calcite, Microcline, Quartz, Gibbsite	Muscovite, Kaolinite, Vermiculite
	40 – 60	Albite, Calcite, Microcline, Quartz, Gibbsite	Muscovite, Kaolinite, Vermiculite
30-Year Cultivated	0 – 20	Albite, Calcite, Microcline, Quartz, Gibbsite	Muscovite, Kaolinite, Vermiculite
	20 – 40	Albite, Chrysotile, Calcite, Microcline, Quartz, Rutile, Gibbsite	Ferruginous Muscovite, Kaolinite, Vermiculite
	40 – 60	Albite, Anatase, Chrysotile, Calcium Marialite, Microcline, Quartz, Gibbsite	Ferruginous Muscovite, Kaolinite, Vermiculite

Appendix H. Results for the Anderson-Darling normality test for data on soil properties for the three intensively monitored plots.

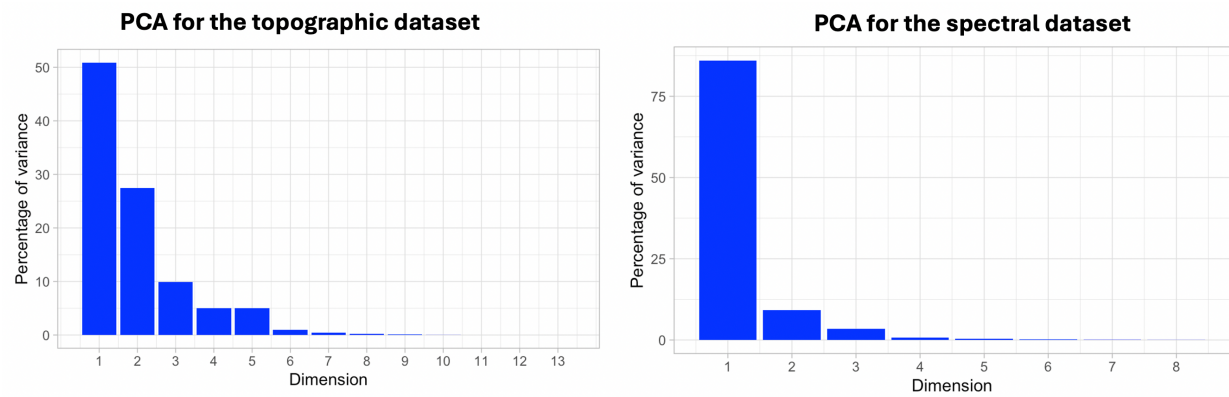
Variable	A statistic	p-value
TN	6.9607	$< 2.2 \times 10^{-16}$
N-NH ₄ ⁺	5.5738	9.083×10^{-14}
N-NO ₃ ⁻	3.1631	6.023×10^{-08}
SOC	5.2211	6.367×10^{-13}
P	26.678	$< 2.2 \times 10^{-16}$
K ⁺	1.5299	6.00×10^{-04}
S	4.5227	3.054×10^{-11}
SWC	3.5914	5.479×10^{-09}
pH	9.9707	$< 2.2 \times 10^{-16}$
BD	5.0421	1.714×10^{-12}
Particle density	4.4684	4.129×10^{-11}
EC	5.6333	6.541×10^{-14}
Total porosity	3.7831	1.878×10^{-09}
Macroporosity	0.47931	0.2235
Mesoporosity	2.2185	1.011×10^{-05}
Microporosity	0.98498	0.01214
Sand	5.8049	2.541×10^{-14}
Silt	8.5664	$< 2.2 \times 10^{-16}$
Clay	3.439	1.285×10^{-08}

Appendix I. Definition and relevance of topographic indexes used for pedotransfer functions development.

Covariate	Definition	Relevance	Reference
Slope	Represents the steepness of the terrain.	Influence flow rates of water and sediment by controlling the rate of energy expenditure or stream power available to drive the flow.	(Zevenbergen and Thorne, 1987)
Aspect	Defines the slope direction (e.g. slope azimuth) and therefore the direction of flow.	Provides the information necessary to determine what upslope land area contributes to the flow at any point in the catchment. Its knowledge also allows analyzing solar radiation, moisture, and wind effects.	(Zevenbergen and Thorne, 1987)
Channel network base level	Elevation relative to the base level of the nearest channel in a drainage network.	Understanding potential erosion or deposition zones.	(Fisher et al., 2017)
Channel network distance	Horizontal distance from a point to the nearest channel.	Mapping hydrological accessibility and drainage patterns.	(Fisher et al., 2017)
Convergence index	Calculates an index of convergence/divergence of overland water flow.	Determining zones of water flow convergence and accumulation.	(Fisher et al., 2017)
Hillshade	Simulates shaded relief based on a light source.	It is used to visualize topography and create maps with a depth effect.	(Fisher et al., 2017)
Length and steepness factor (LS-Factor)	Is a combination of Slope and Slope length which combined are key attribute for predicting erosion potential within a landscape.	Predicting erosion-prone areas for management planning.	(Fisher et al., 2017)
Plan curvature	Curvature of the land surface transverse to the slope direction.	Influences flow convergence and divergence. Can be useful for modelling landslide risk.	(Fisher et al., 2017)
Profile curvature	The rate of change of slope.	Affects flow acceleration and deceleration and therefore influences aggradation and degradation. Can be useful for modelling landslide risk.	(Fisher et al., 2017; Zevenbergen and Thorne, 1987)

Covariate	Definition	Relevance	Reference
Relative slope position	Relative position on a slope between its top and bottom.	Modeling runoff, erosion, and suitability for agriculture.	(Fisher et al., 2017)
Topographic wetness index	Produces a grid showing water accumulation.	Can be useful for soil or flood mapping and predicting soil moisture and hydrological processes.	(Fisher et al., 2017)
Total catchment area	Total area upstream contributing surface runoff to a point.	Allows to identify areas of soil moisture accumulation derived from runoff.	(Moore et al., 1991)
Valley depth	Depth relative to the highest point in surrounding ridgelines.	Identifying valley floors and erosion patterns.	(Fisher et al., 2017)

Appendix J. Decomposition of the total inertia for the principal component analyses for groups of remote sensing covariates.



Appendix K. Comparison of the probability density functions of the three groups of covariates.

Figure K1. Comparison of the probability density functions of the primary and secondary edaphic properties datasets.

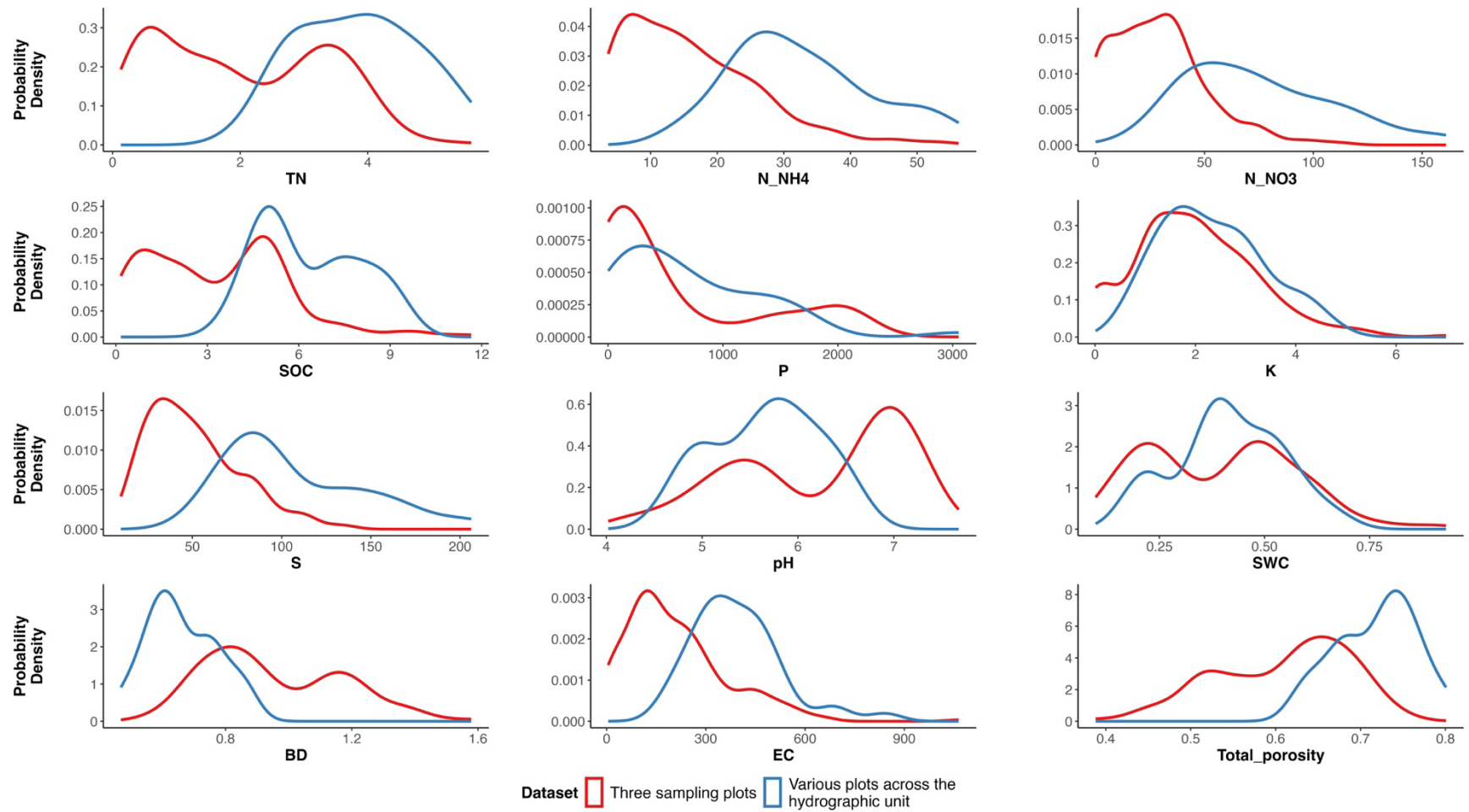


Figure K2. Comparison of the probability density functions of the primary and secondary topographic indexes datasets.

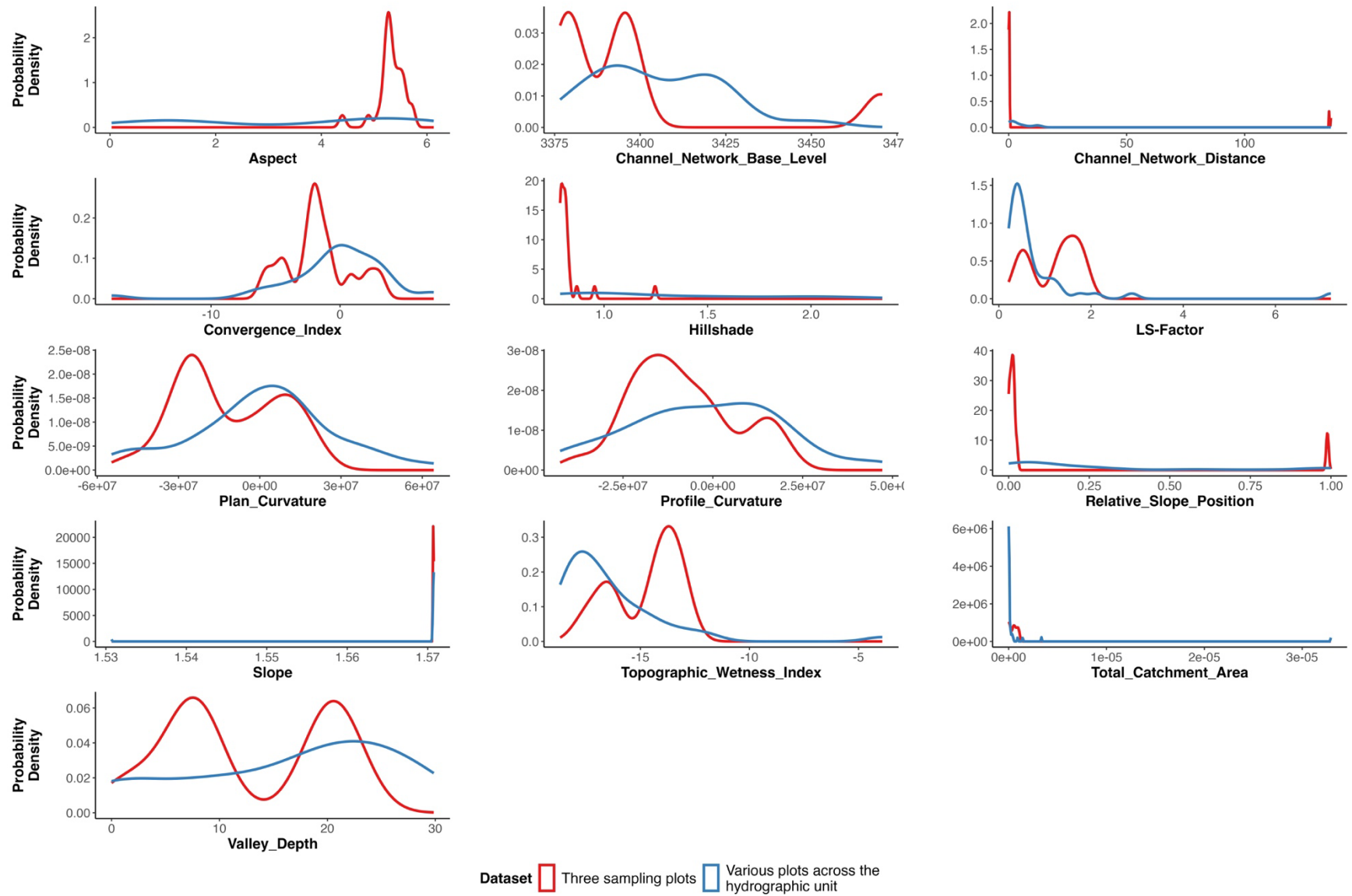
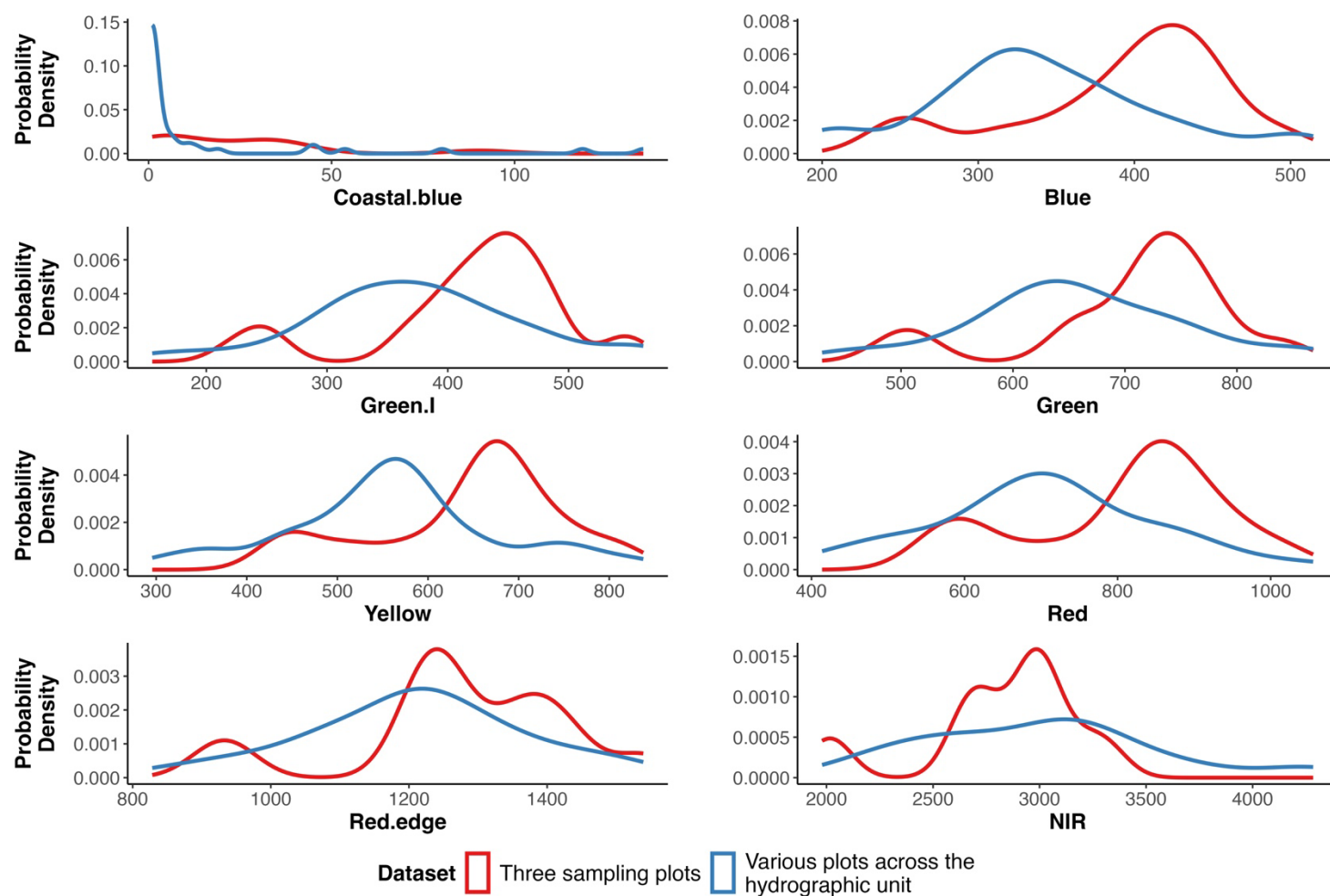


Figure K3. Comparison of the probability density functions of the primary and secondary spectral signature datasets.

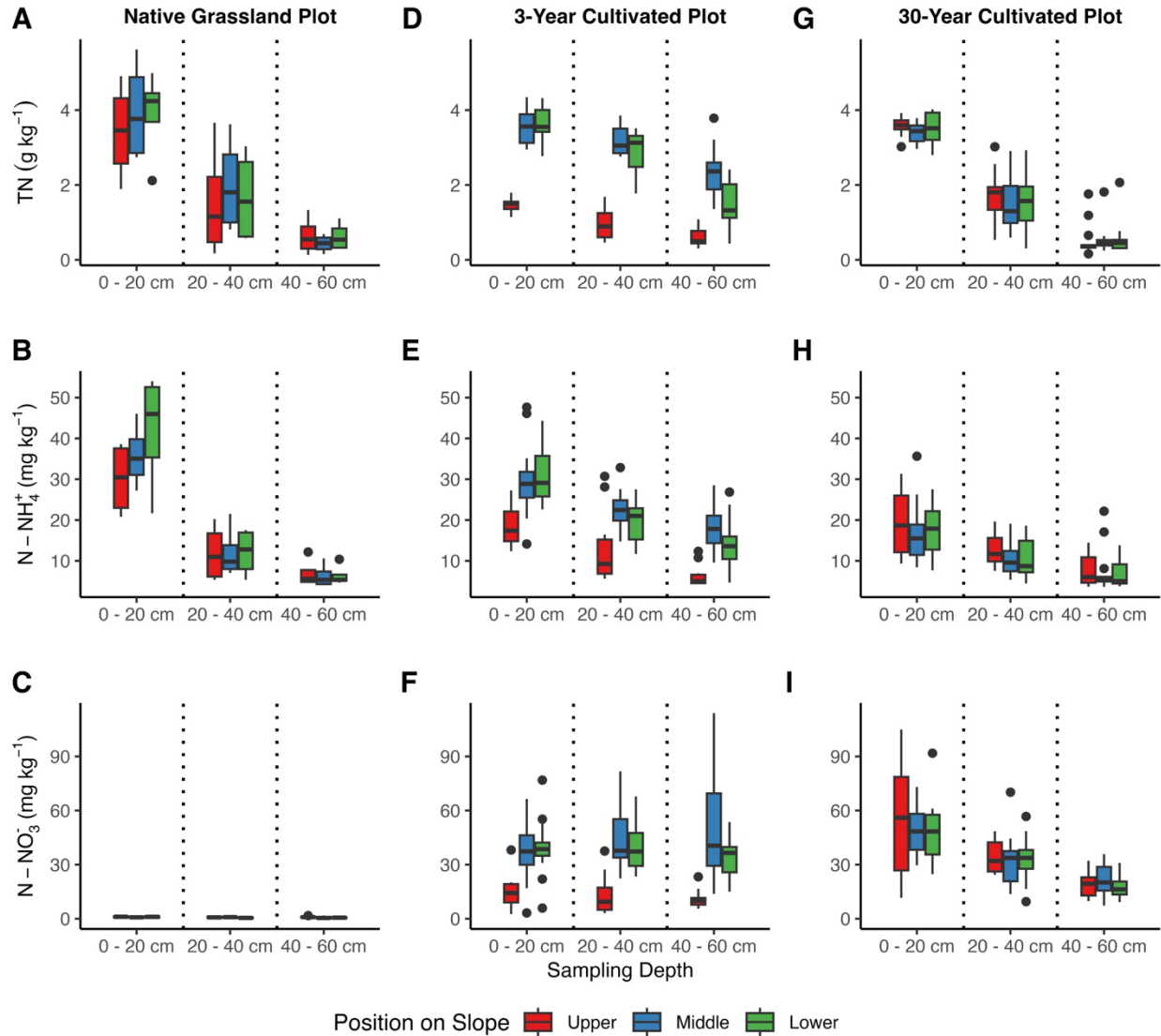


Appendix L. Pearson correlation coefficients (r) between the data of the original variables (i.e., edaphic properties) and the scores of the first three principal components for data on the three intensively monitored plots.

Variable	Principal component 1	Principal component 2	Principal component 3
TN	0.89**	0.09	0.27**
N-NH ₄ ⁺	0.71**	-0.20**	0.26**
N-NO ₃ ⁻	0.61**	0.51**	-0.22**
SOC	0.85**	-0.20**	0.30**
P	0.21**	0.84**	0.32**
K ⁺	0.43**	0.75**	0.07
S	0.56**	-0.12*	-0.52**
pH	-0.48**	0.77**	0.11
SWC	0.86**	-0.28**	0.18**
BD	-0.86**	0.18**	-0.27**
Particle density	-0.86**	0.12*	0.04
EC	0.58**	0.72**	0.07
Total porosity	0.81**	-0.19**	0.31**
Sand	-0.83**	-0.06	0.48**
Silt	0.84**	-0.08	-0.23**
Clay	0.53**	0.22**	-0.66**

Note. Significant correlations are indicated * $p < 0.05$; ** $p < 0.01$.

Appendix M. Variation in soil TN, N-NH₄⁺, and N-NO₃⁻ concentrations across three sampling plots and soil depths with data grouped by position on slope. Each column represents data from one plot: native grassland plot (A–C), 3-Year cultivated plot (D–F), and 30-Year cultivated plot (G–I).



Appendix N. Spearman correlations coefficients (ρ) among soil N and measured edaphic properties using data for the three intensively monitored plots.

Variable	TN	N-NH ₄ ⁺	N-NO ₃ ⁻
N-NH ₄ ⁺	0.783**		
N-NO ₃ ⁻	0.529**	0.371**	
SOC	0.916**	0.802**	0.453**
P	0.225**	0.002	0.520**
K	0.521**	0.368**	0.590**
S	0.474**	0.363**	0.496**
pH	-0.389**	-0.407**	0.007
SWC	0.759**	0.665**	0.370**
BD	-0.724**	-0.624**	-0.407**
Particle density	-0.690**	-0.637**	-0.513**
EC	0.667**	0.447**	0.785**
Total porosity	0.691**	0.580**	0.363**
Macroporosity	-0.428**	-0.355*	-0.321*
Mesoporosity	-0.373*	-0.302*	-0.194
Microporosity	0.794**	0.699**	0.504**
Sand	-0.628**	-0.562**	-0.592**
Silt	0.660**	0.611**	0.507**
Clay	0.383**	0.308**	0.519**

Note. Significant correlations are indicated * $p < 0.05$; ** $p < 0.01$.

Appendix O. Full set of the results of the Linear mixed-effects models

Table O1. Linear mixed-effects model results for total nitrogen

	Core LMM		3-year cultivated LMM		30-year cultivated LMM	
Random effects	Variance (SD)		Variance (SD)		Variance (SD)	
Sampling point	0.41 (0.64)		0.95 (0.98)		0.025 (0.15)	
Residual	0.31 (0.56)		0.21 (0.45)		0.19 (0.43)	
Fixed effects	Estimate (CI 95%)	p-value	Estimate (CI 95%)	p-value	Estimate (CI 95%)	p-value
Intercept ^a	3.9*** (3.14 – 4.66)	0.000	2.48 (1.58 – 3.37)	0.001	3.53*** (3.19 – 3.87)	0.000
Plot (Land-use)						
3-year cultivated	-0.97 (-1.88 – -0.06)	0.066				
30-year cultivated	-0.37 (-1.28 – 0.54)	0.461				
Depth						
20 - 40 cm	-2.02*** (-2.46 – -1.58)	0.000	-0.43 (-0.91 – 0.04)	0.101	-1.76*** (-2.22 – -1.31)	0.000
40 - 60 cm	-3.2*** (-3.64 – -2.76)	0.000	-1.38*** (-1.86 – -0.91)	0.000	-2.91*** (-3.36 – -2.46)	0.000
Month						
2nd	-0.22 (-0.45 – 0.01)	0.066	0.65* (0.18 – 1.13)	0.014	-0.12 (-0.57 – 0.33)	0.638
3rd	-0.29* (-0.52 – -0.07)	0.014	0.14 (-0.34 – 0.61)	0.605	-0.09 (-0.54 – 0.37)	0.735
4th	0.02 (-0.22 – 0.27)	0.867	0.36 (-0.11 – 0.84)	0.168	-0.06 (-0.51 – 0.4)	0.823
5th	0.09 (-0.16 – 0.33)	0.502	0.69** (0.22 – 1.17)	0.010	-0.06 (-0.52 – 0.39)	0.799
6th	-0.02 (-0.25 – 0.21)	0.861	0.41 (-0.06 – 0.88)	0.120	-0.13 (-0.59 – 0.32)	0.595
7th	0.18 (-0.07 – 0.42)	0.163	0.67* (0.2 – 1.14)	0.012	0.19 (-0.26 – 0.65)	0.442
Plot (land-use) x Depth						
3-year cultivated x 20 - 40 cm	1.46*** (0.96 – 1.96)	0.000				
30-year cultivated x 20 - 40 cm	0.09 (-0.41 – 0.59)	0.732				
3-year cultivated x 40 - 60 cm	1.76*** (1.26 – 2.26)	0.000				

	Core LMM	3-year cultivated LMM	30-year cultivated LMM
Random effects	Variance (SD)	Variance (SD)	Variance (SD)
30-year cultivated x 40 - 60 cm	0.24 (-0.26 – 0.74)	0.348	
Depth x Month			
20 - 40 cm x 2nd month		-0.06 (-0.73 – 0.61) 0.864	-0.71* (-1.35 – -0.07) 0.050
40 - 60 cm x 2nd month		-0.19 (-0.86 – 0.48) 0.608	-0.1 (-0.74 – 0.54) 0.781
20 - 40 cm x 3rd month		-0.01 (-0.67 – 0.66) 0.989	0 (-0.64 – 0.64) 0.990
40 - 60 cm x 3rd month		0.26 (-0.41 – 0.93) 0.489	-0.08 (-0.72 – 0.56) 0.814
20 - 40 cm x 4th month		-0.19 (-0.86 – 0.48) 0.612	0.24 (-0.4 – 0.88) 0.494
40 - 60 cm x 4th month		0.06 (-0.61 – 0.73) 0.874	0.05 (-0.59 – 0.69) 0.899
20 - 40 cm x 5th month		-0.13 (-0.8 – 0.54) 0.722	-0.33 (-0.97 – 0.31) 0.356
40 - 60 cm x 5th month		0.01 (-0.66 – 0.68) 0.987	0.04 (-0.6 – 0.68) 0.918
20 - 40 cm x 6th month		-0.17 (-0.84 – 0.5) 0.651	-0.47 (-1.11 – 0.17) 0.184
40 - 60 cm x 6th month		-0.37 (-1.04 – 0.3) 0.317	-0.06 (-0.7 – 0.58) 0.874
20 - 40 cm x 7th month		-0.32 (-0.99 – 0.35) 0.395	0.08 (-0.56 – 0.72) 0.828
40 - 60 cm x 7th month		-0.15 (-0.81 – 0.52) 0.695	-0.18 (-0.82 – 0.46) 0.618

Table O2. Linear mixed-effects model results for ammonium nitrogen

	Core LMM		3-year cultivated LMM		30-year cultivated LMM	
Random effects	Variance (SD)		Variance (SD)		Variance (SD)	
Sampling point	12.18 (3.49)		28.32 (5.32)		0.46 (0.68)	
Residual	27.55 (5.25)		30.64 (5.54)		10.75 (3.28)	
Fixed effects	Estimate (CI 95%)	p-value	Estimate (CI 95%)	p-value	Estimate (CI 95%)	p-value
Intercept ^a	36.79*** (31.93 – 41.65)	0.000	27.65*** (21.73 – 33.57)	0.000	24.02*** (21.56 – 26.49)	0.000
Plot (Land-use)						
3-year cultivated	-10.46** (-16.09 – -4.83)	0.002				
30-year cultivated	-19.08*** (-24.71 – -13.45)	0.000				
Depth						
20 - 40 cm	-23.91*** (-28.03 – -19.79)	0.000	-11.67*** (-17.43 – -5.9)	0.000	-12.65*** (-16.06 – -9.23)	0.000
40 - 60 cm	-29.33*** (-33.45 – -25.21)	0.000	-19.53*** (-25.3 – -13.77)	0.000	-18.7*** (-22.11 – -15.28)	0.000
Month						
2nd	-1.36 (-3.48 – 0.77)	0.221	0.29 (-5.47 – 6.06)	0.927	-12.8*** (-16.22 – -9.39)	0.000
3rd	-3.87*** (-6 – -1.74)	0.001	-7.09* (-12.86 – -1.33)	0.029	-11.19*** (-14.6 – -7.77)	0.000
4th	-2.15 (-4.44 – 0.13)	0.071	-3.59 (-9.35 – 2.17)	0.264	-13.41*** (-16.83 – -10)	0.000
5th	0.92 (-1.37 – 3.21)	0.441	4.12 (-1.65 – 9.88)	0.201	-3.87* (-7.28 – -0.45)	0.044
6th	1.81 (-0.32 – 3.94)	0.103	-4.17 (-9.93 – 1.6)	0.195	-2.84 (-6.25 – 0.58)	0.137
7th	5.49*** (3.2 – 7.77)	0.000	2.04 (-3.73 – 7.8)	0.525	0.77 (-2.64 – 4.19)	0.683
Plot (land-use) x Depth						
3-year cultivated x 20 - 40 cm	15.65*** (10.98 – 20.32)	0.000				
30-year cultivated x 20 - 40 cm	17.28*** (12.61 – 21.95)	0.000				
3-year cultivated x 40 - 60 cm	15.47*** (10.8 – 20.14)	0.000				
30-year cultivated x 40 - 60 cm	18.8*** (14.12 – 23.47)	0.000				
Depth x Month						

	Core LMM	3-year cultivated LMM	30-year cultivated LMM	
Random effects	Variance (SD)	Variance (SD)	Variance (SD)	
20 - 40 cm x 2nd month		5.72 (-2.44 – 13.87)	0.209	9.45*** (4.62 – 14.27) 0.001
40 - 60 cm x 2nd month		4.33 (-3.82 – 12.48)	0.340	12.44*** (7.61 – 17.27) 0.000
20 - 40 cm x 3rd month		6.27 (-1.88 – 14.42)	0.168	9.44*** (4.62 – 14.27) 0.001
40 - 60 cm x 3rd month		10.21* (2.06 – 18.37)	0.026	10.67*** (5.84 – 15.5) 0.000
20 - 40 cm x 4th month		5.29 (-2.86 – 13.45)	0.244	12.14*** (7.31 – 16.97) 0.000
40 - 60 cm x 4th month		7.11 (-1.04 – 15.26)	0.119	13.89*** (9.06 – 18.72) 0.000
20 - 40 cm x 5th month		-0.87 (-9.02 – 7.29)	0.848	1.5 (-3.33 – 6.33) 0.576
40 - 60 cm x 5th month		0.43 (-7.72 – 8.58)	0.924	4.06 (-0.77 – 8.89) 0.133
20 - 40 cm x 6th month		6.34 (-1.81 – 14.5)	0.164	5.06 (0.23 – 9.89) 0.062
40 - 60 cm x 6th month		7.6 (-0.55 – 15.75)	0.096	6.78* (1.95 – 11.61) 0.013
20 - 40 cm x 7th month		1.05 (-7.1 – 9.2)	0.817	4.5 (-0.33 – 9.33) 0.096
40 - 60 cm x 7th month		10.0* (1.85 – 18.15)	0.029	9.29*** (4.46 – 14.12) 0.001

Table O3. Linear mixed-effects model results for nitrate nitrogen

	Core LMM		3-year cultivated LMM		30-year cultivated LMM	
Random effects	Variance (SD)		Variance (SD)		Variance (SD)	
Sampling point	95.55 (9.78)		234.1 (15.30)		0.0 (0.0)	
Residual	171.84 (13.11)		162.20		127.20 (11.28)	
Fixed effects	Estimate (CI 95%)	p-value	Estimate (CI 95%)	p-value	Estimate (CI 95%)	p-value
Intercept ^a	8.11 (-4.94 – 21.15)	0.258	35.52*** (19.83 – 51.21)	0.001	82.05*** (73.75 – 90.35)	0.000
Plot (Land-use)						
3-year cultivated	30.88** (15.66 – 46.1)	0.001				
30-year cultivated	51.01*** (35.79 – 66.24)	0.000				
Depth						
20 - 40 cm	-0.21 (-10.5 – 10.08)	0.968	11.09 (-2.18 – 24.35)	0.135	-47.82*** (-59.56 – -36.08)	0.000
40 - 60 cm	-0.16 (-10.45 – 10.13)	0.976	9.25 (-4.02 – 22.51)	0.212	-66.79*** (-78.53 – -55.05)	0.000
Month						
2nd	-2.83 (-8.14 – 2.49)	0.307	11.09 (-2.17 – 24.36)	0.135	-16.18* (-27.92 – -4.44)	0.015
3rd	-10.97*** (-16.28 – -5.65)	0.000	-12.62 (-25.88 – 0.65)	0.089	-41.92*** (-53.66 – -30.17)	0.000
4th	-11.23*** (-16.94 – -5.52)	0.000	-11.03 (-24.3 – 2.23)	0.137	-28.06*** (-39.8 – -16.32)	0.000
5th	-15.35*** (-21.06 – -9.64)	0.000	-19.14* (-32.4 – -5.87)	0.011	-51.0*** (-62.74 – -39.26)	0.000
6th	-15.07*** (-20.38 – -9.76)	0.000	-7.98 (-21.24 – 5.28)	0.280	-53.75*** (-65.49 – -42.01)	0.000
7th	-2.44 (-8.16 – 3.27)	0.412	6.05 (-7.22 – 19.31)	0.413	-27.5*** (-39.24 – -15.75)	0.000
Plot (land-use) x Depth						
3-year cultivated x 20 - 40 cm	1.85 (-9.81 – 13.52)	0.760				
30-year cultivated x 20 - 40 cm	-17.35** (-29.02 – -5.68)	0.005				
3-year cultivated x 40 - 60 cm	0.54 (-11.13 – 12.21)	0.929				
30-year cultivated x 40 - 60 cm	-31.25*** (-42.91 – -19.58)	0.000				

	Core LMM	3-year cultivated LMM	30-year cultivated LMM
Random effects	Variance (SD)	Variance (SD)	Variance (SD)
Depth x Month			
20 - 40 cm x 2nd month		-21.62* (-40.38 – -2.86) 0.040	9.56 (-7.04 – 26.16) 0.302
40 - 60 cm x 2nd month		-10.91 (-29.67 – 7.84) 0.297	16.52 (-0.08 – 33.13) 0.076
20 - 40 cm x 3rd month		-0.98 (-19.73 – 17.78) 0.925	41.09*** (24.49 – 57.69) 0.000
40 - 60 cm x 3rd month		-3.07 (-21.83 – 15.69) 0.769	44.76*** (28.16 – 61.36) 0.000
20 - 40 cm x 4th month		-13.06 (-31.81 – 5.7) 0.212	31.96*** (15.36 – 48.56) 0.001
40 - 60 cm x 4th month		-13.37 (-32.13 – 5.39) 0.202	33.3*** (16.7 – 49.91) 0.000
20 - 40 cm x 5th month		2.86 (-15.89 – 21.62) 0.784	50.03*** (33.43 – 66.63) 0.000
40 - 60 cm x 5th month		0.42 (-18.34 – 19.17) 0.968	53.94*** (37.33 – 70.54) 0.000
20 - 40 cm x 6th month		-17.23 (-35.99 – 1.53) 0.101	47.45*** (30.85 – 64.06) 0.000
40 - 60 cm x 6th month		-18.68 (-37.43 – 0.08) 0.076	59.74*** (43.13 – 76.34) 0.000
20 - 40 cm x 7th month		-16.11 (-34.86 – 2.65) 0.125	31.72*** (15.11 – 48.32) 0.001
40 - 60 cm x 7th month		-16.44 (-35.2 – 2.32) 0.117	39.46*** (22.86 – 56.07) 0.000

Appendix P. Results of the Recursive Feature Elimination

Figure P1. Frequency of covariate selection for predicting soil TN (A), N-NH₄⁺ (B), and N-NO₃⁻ (C) using recursive feature elimination for the S1 scenario, with data for three sampling depths.

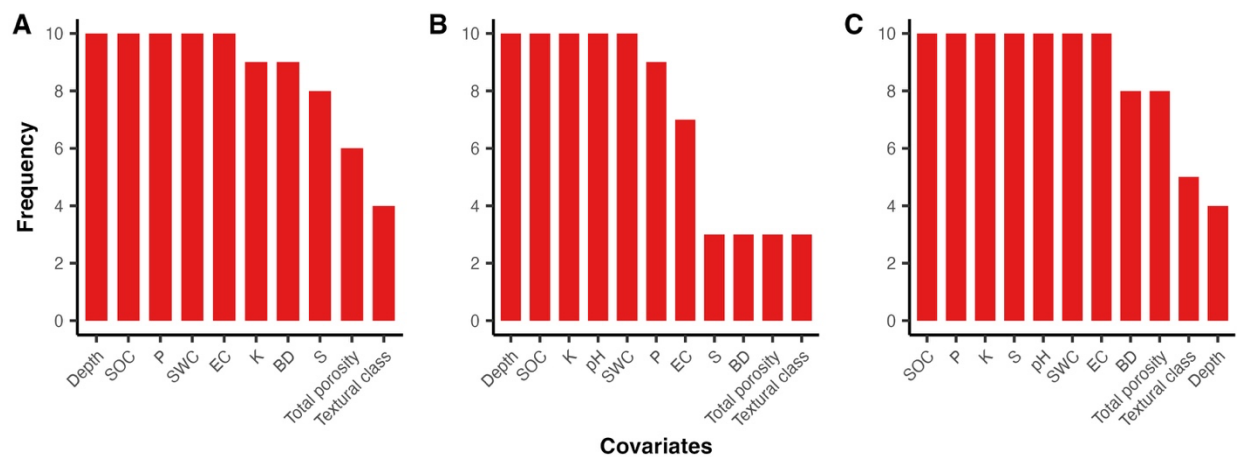
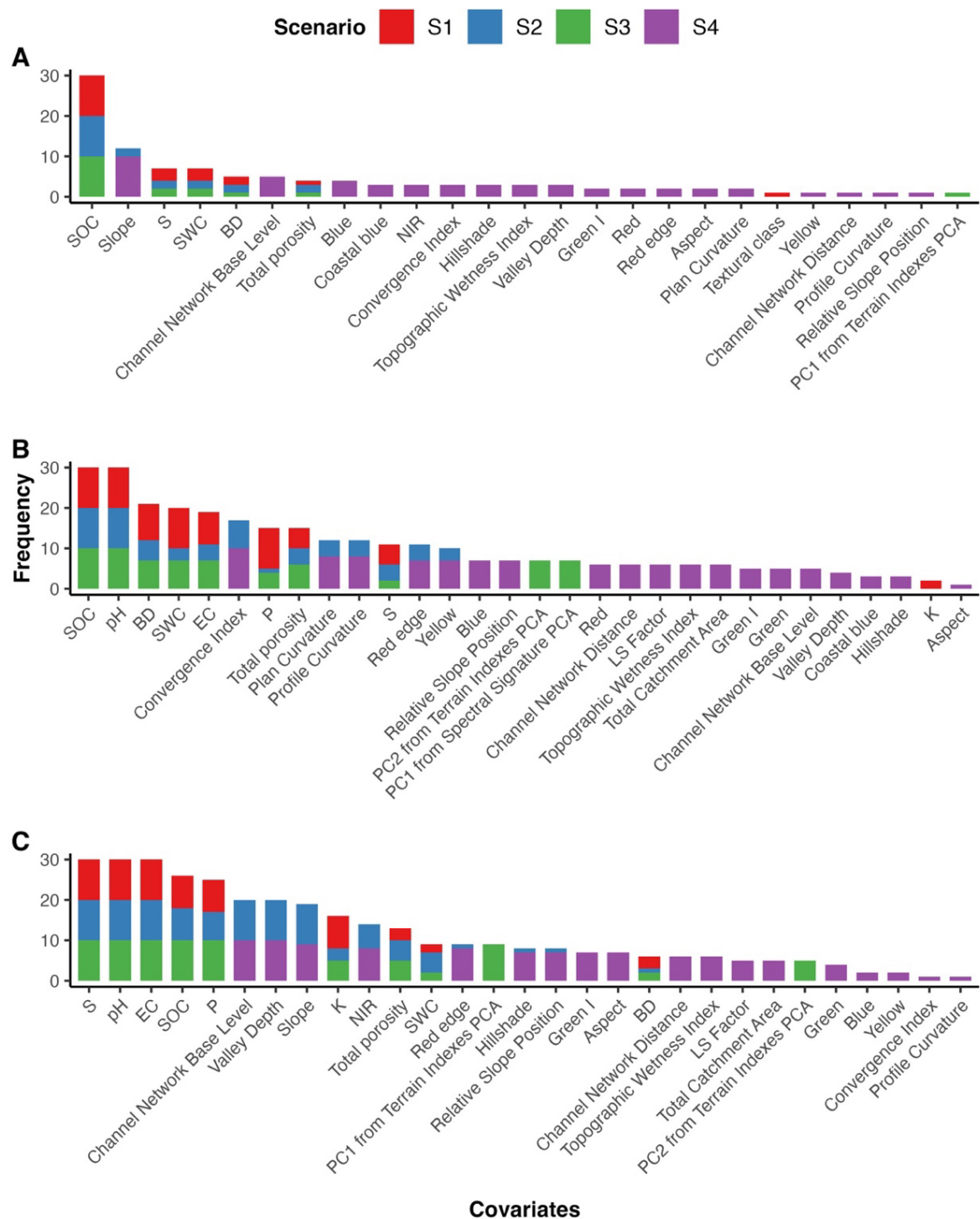
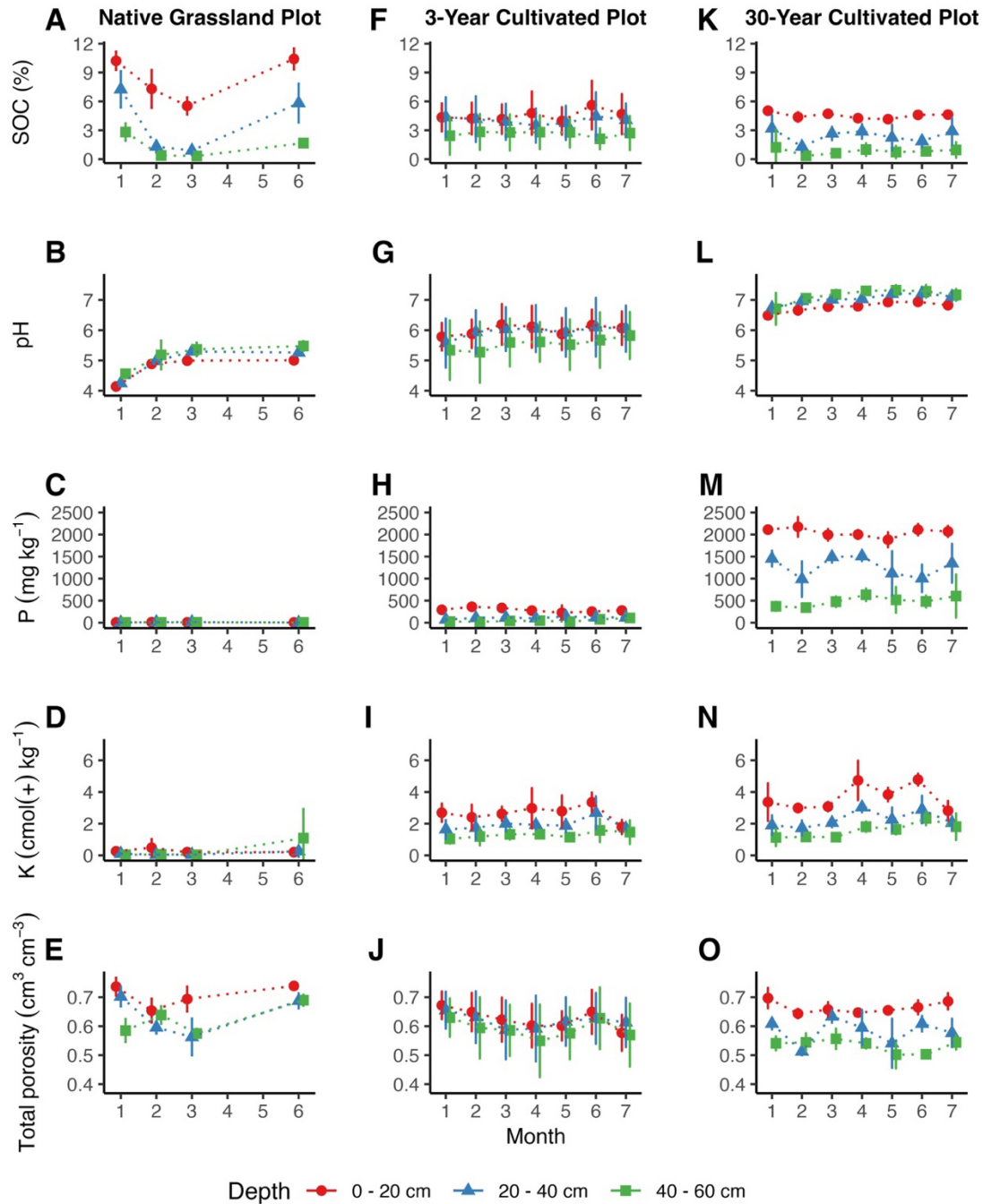


Figure P2. Frequency of covariate selection for predicting soil TN (A), N-NH₄⁺ (B), and N-NO₃⁻ (C) using recursive feature elimination across four modeling scenarios, with data for 0-20 cm sampling depth.



Appendix Q. Temporal variation of edaphic properties across the monitoring period and the three sampling plots: Native grassland plot (A–E), 3-Year cultivated plot (F–J), and 30-Year cultivated plot (K–O). Points represent mean concentrations, with error bars indicating standard deviations.



Appendix R. Statistical analyses for pedotransfer function development and evaluation.

Table R1. Summary of linear models for predicting soil N pools.

Covariates	Estimate	Standard error	t value	p-value
<i>Multi-depth PTF por predicting soil TN</i>				
Intercept	3.576 E+00	2.179 E+00	1.641	0.102
SOC	3.921 E-01	1.964 E-02	19.965	0.000
`Depth20 - 40 cm`	-2.643 E-01	6.576 E-02	-4.018	0.000
`Depth40 - 60 cm`	-3.748 E-01	8.585 E-02	-4.365	0.000
EC	1.447 E-03	2.976 E-04	4.864	0.000
P	8.273 E-05	5.180 E-05	1.597	0.111
SWC	4.635 E-01	2.818 E-01	1.645	0.101
BD	-1.318 E+00	8.451 E-01	-1.559	0.120
K	7.795 E-02	2.753 E-02	2.832	0.005
S	1.941 E-03	1.094 E-03	1.775	0.077
Total_porosity	-3.547 E+00	2.192 E+00	-1.618	0.107
<i>Topsoil PTF por predicting soil TN</i>				
Intercept	1.60714	0.18544	8.667	0.000
SOC	0.33209	0.03388	9.801	0.000
<i>Topsoil PTF por predicting soil N-NO₃⁻</i>				
Intercept	-6.7323	37.1321	-0.181	0.857
EC	0.0551	0.0148	3.722	0.000
S	0.6637	0.1046	6.344	0.000
pH	-0.7583	5.4033	-0.140	0.889
P	0.0072	0.0036	1.989	0.050
SOC	-1.5592	1.3836	-1.127	0.263
K	-1.4714	1.6476	-0.893	0.375

Figure R1. PCA factor map for the covariates used in TN multi-depth PTF.

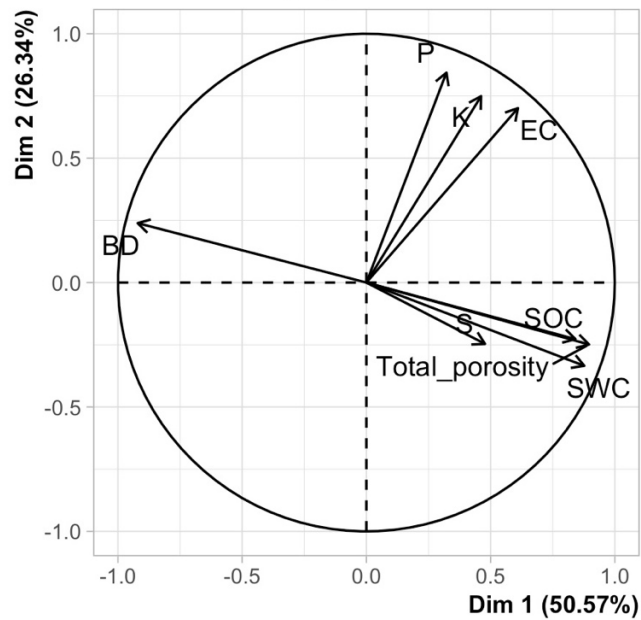
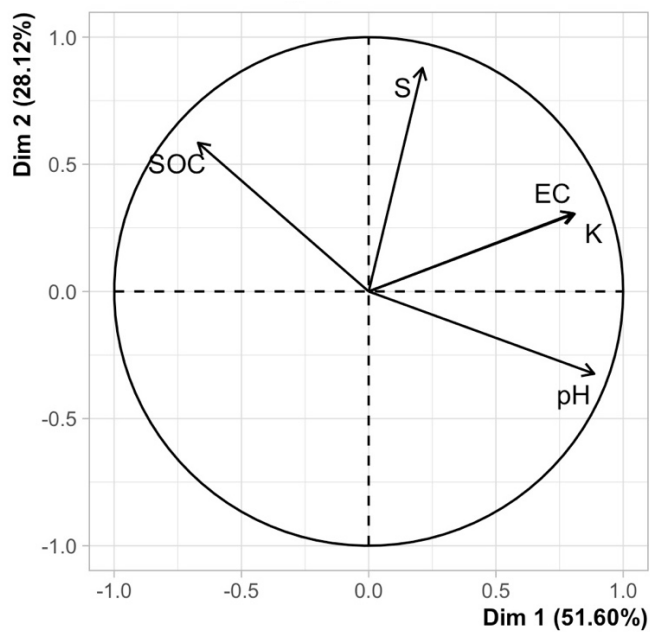


Figure R2. PCA factor map for the covariates used in N-NO₃⁻ topsoil PTF. Only predictors explaining high proportion of variance are shown.



Appendix S. Purity and concentration of genomic DNA extracted from soil samples

Sample ID	Downstream analysis*	Concentration (ng μL^{-1})	A260/A280	A260/A230
1-N1-A	A	221.8	1.80	2.00
	B	223.7	1.92	0.39
1-N2-A	A	238.5	1.81	2.06
	B	297.6	1.92	0.99
1-N3-A	A	210.4	1.78	1.82
	B	260.0	1.90	1.40
1-C1-A	A	80.7	1.83	1.91
	B	142.7	1.91	0.33
1-C3-A	A	105.4	1.84	1.96
	B	191.2	1.91	0.94
1-C5-A	A	132.3	1.83	1.74
	B	96.5	1.91	1.62
1-C7-A	A	166.3	1.83	1.64
	B	202.2	1.91	1.07
1-C9-A	A	129.4	1.80	1.94
	B	121.8	1.89	0.97
1-C11-A	A	158.4	1.81	1.70
	B	161.5	1.90	1.59
2-N1-A	A	99.0	1.93	1.70
	B	58.7	1.93	0.22
2-N2-A	A	233.0	1.93	1.38
	B	249.6	1.91	1.02
2-N3-A	A	300.0	1.92	1.95
	B	263.7	1.92	0.88
2-C1-A	A	89.6	1.91	1.60
	B	99.1	1.92	0.97
2-C3-A	A	162.1	1.91	1.62
	B	162.9	1.91	0.48
2-C5-A	A	131.2	1.91	2.17
	B	94.15	1.92	0.72
2-C7-A	A	203.1	1.91	1.61
	B	174.3	1.87	1.58
2-C9-A	A	212.2	1.91	2.23
	B	161.4	1.90	1.76
2-C11-A	A	143.4	1.80	1.92
	B	94.6	1.75	0.48

Sample ID	Downstream analysis*	Concentration (ng μL^{-1})	A260/A280	A260/A230
3-N1-A	A	168.6	1.82	2.00
	B	210.2	1.83	1.40
3-N2-A	A	252.8	1.87	1.93
	B	286.7	1.88	1.32
3-N3-A	A	187.7	1.84	1.73
	B	224.2	1.86	1.56
3-C1-A	A	122.5	1.77	1.75
	B	123.2	1.73	1.64
3-C3-A	A	145.6	1.80	1.38
	B	169.4	1.78	1.00
3-C5-A	A	89.6	1.73	1.77
	B	124.5	1.74	0.93
3-C7-A	A	109.1	1.72	1.74
	B	174.3	1.79	1.31
3-C9-A	A	207.8	1.72	1.58
	B	170.7	1.82	0.45
3-C11-A	A	202.2	1.80	1.74
	B	194.0	1.79	0.60

*A: high-throughput sequencing, B: qPCR. The sample ID coding is structured as follows: the first number indicates the sampling month; the following letter denotes the land use type (N: native, C: cultivated); the next number identifies the position of the sampling point within the plot; and the final letter specifies that the soil sample was collected from the surface layer (A: 0–20 cm depth).

Appendix T. Summary of quality assessment of high throughput DNA sequence data

MiSeq Run	Sample ID	Read	Position in read where quality index (phred score) drops under 30			
			Mean	Median	Q1	Min
September 2023	1-N1-A	R1	300	300	285	225
		R2	225	240	200	125
	1-N2-A	R1	300	300	279	225
		R2	225	250	200	125
	1-N3-A	R1	300	300	275	225
		R2	225	225	200	125
	1-C1-A	R1	300	300	279	240
		R2	320	240	200	125
	1-C3-A	R1	300	300	279	225
		R2	225	250	200	125
	1-C5-A	R1	279	279	250	175
		R2	179	179	130	7
	1-C7-A	R1	300	300	275	200
		R2	220	225	175	64
	1-C9-A	R1	300	300	279	225
		R2	225	240	200	125
	1-C11-A	R1	300	300	279	200
		R2	225	240	200	125
January 2024	2-N1-A	R1	210	229	200	154
		R2	250	260	240	204
	2-N2-A	R1	204	225	179	150
		R2	250	279	230	200
	2-N3-A	R1	200	225	175	150
		R2	250	254	225	200
	2-C1-A	R1	225	230	185	150
		R2	250	275	250	200
	2-C3-A	R1	225	250	200	154
		R2	250	275	250	200
	2-C5-A	R1	225	240	200	150
		R2	250	275	230	200
	2-C7-A	R1	225	229	200	175
		R2	250	275	240	200
	2-C9-A	R1	225	240	200	160
		R2	250	275	250	200

MiSeq Run	Sample ID	Read	Position in read where quality index (phred score) drops under 30			
			Mean	Median	Q1	Min
	2-C11-A	R1	225	250	200	175
		R2	250	240	250	200
	3-N1-A	R1	210	229	179	150
		R2	250	260	240	200
	3-N2-A	R1	210	240	200	150
		R2	250	275	240	200
	3-N3-A	R1	200	225	175	150
		R2	250	254	225	200
	3-C1-A	R1	229	250	200	175
		R2	254	275	250	200
	3-C3-A	R1	225	250	200	150
		R2	270	275	250	200
	3-C5-A	R1	225	250	200	175
		R2	254	275	240	200
	3-C7-A	R1	225	240	200	150
		R2	254	275	250	200
	3-C9-A	R1	225	240	200	154
		R2	250	260	240	200
	3-C11-A	R1	210	229	200	150
		R2	254	275	250	200

The sample ID coding is structured as follows: the first number indicates the sampling month; the following letter denotes the land use type (N: native, C: cultivated); the next number identifies the position of the sampling point within the plot; and the final letter specifies that the soil sample was collected from the surface layer (A: 0–20 cm depth).

Appendix U. Quality filtering criteria and final DADA2 parameters used for sequence processing

MiSeq Run	Filtering criteria parameters for the DADA2 step				Percentage of input passed filter	Sequence Count
	trim-left-f	trim-left-r	trunc-len-f	trunc-len-r		
Sep 2023	30	30	280	200	Most samples (n=8): 78-83% Sample 1-C5-A: 62%	7620
Jan 2024	30	30	210	270	All samples (n=18): 80-85%	5259

Appendix V. Primer sequences, annealing temperatures, and qPCR conditions used for the quantification of nitrogen-cycling genes

Target gene	Primers	References	Amplicon length (bp)	Annealing temperature (°C)	Efficiency (%)	Standard curve r^2
	Sequence (5'→3')					
<i>nifH</i>	Forward: AAAGGYGGWATCGGYAARTCCACCAC Reverse: TTGTTSGCSGCRTACATSGCCATCA	(Rösch et al., 2002)	461	56	75.4	0.996
<i>nosZ</i>	Forward: CGYTGTTCMTGACAGCCAG Reverse: CGSACCTTSTTGCCSTYGCG	(Rösch et al., 2002; Throbäck et al., 2004)	598	54	128.9	0.968
<i>amoA</i> (AOB)	Forward: GGGGHTTYTACTGGTGGT Reverse: CCCCTCNGNAAAGCCTTCTTC	(Laverman et al., 2001; Rotthauwe et al., 1997)	491	55	105.3	0.986
<i>amoA</i> (AOA)	Forward: TAATGGTCTGGCTTAGACG Reverse: CGGCCATCCATCTGTATGT	(Liu et al., 2024)	633	55	99	0.990
<i>nirS</i>	Forward: AACGYSAAGGARACSGG Reverse: GASTTCGGRTGSGTCTTSAYGAA	(Levy-Booth et al., 2014)	425	54	131.3	0.961

Appendix W. Summary of statistical comparisons of dominant phyla abundances using Welch's t-test with Bonferroni correction to evaluate differences between native grassland and cultivated plots.

Phylum	Cultivated plots		Native grassland plot		p-value	Corrected p-value	Difference between means
	Mean rel. freq. (%)	Std. dev. (%)	Mean rel. freq. (%)	Std. dev. (%)			
Planctomycetota	6.108	1.306	9.082	1.409	0.000	0.007	-2.974
Chloroflexi	5.495	1.488	9.134	1.485	0.000	0.002	-3.639
Proteobacteria	28.391	3.158	39.261	3.840	0.000	0.000	-10.869
Bacteroidota	5.873	1.960	1.668	1.345	0.000	0.000	4.205
Myxococcota	2.403	0.457	1.408	0.278	0.000	0.000	0.995
Verrucomicrobiota	4.598	1.159	10.223	1.308	0.000	0.000	-5.624
Actinobacteriota	19.567	2.277	9.047	2.056	0.000	0.000	10.521
Acidobacteriota	8.252	1.240	17.054	1.379	0.000	0.000	-8.802
Gemmatimonadota	1.539	0.335	0.206	0.082	0.000	0.000	1.333
Firmicutes	13.595	2.561	0.941	0.715	0.000	0.000	12.654

Appendix X. Summary of statistical comparisons of dominant phyla abundances using Welch's t-test with Bonferroni correction to evaluate differences between the cultivated plots.

Phylum	3-year cultivated plot		30-year cultivated plot		p-value	Corrected p-value	Difference between means
	Mean rel. freq. (%)	Std. dev. (%)	Mean rel. freq. (%)	Std. dev. (%)			
Gemmatimonadota	1.834	0.171	1.244	0.145	0.000	0.000	0.590
Chloroflexi	4.338	0.618	6.651	1.171	0.000	0.015	-2.312
Verrucomicrobiota	5.243	0.866	3.954	1.052	0.017	0.782	1.289
Myxococcota	2.570	0.473	2.237	0.372	0.139	6.383	0.333
Actinobacteriota	20.301	2.327	18.834	1.968	0.193	8.865	1.467
Bacteroidota	6.383	2.247	5.362	1.455	0.299	13.765	1.021
Acidobacteriota	8.037	1.400	8.467	1.011	0.493	22.665	-0.430
Proteobacteria	28.100	0.857	28.682	4.364	0.720	33.122	-0.582
Planctomycetota	6.224	1.149	5.992	1.437	0.727	33.432	0.232
Firmicutes	13.376	2.347	13.813	2.741	0.737	33.883	-0.437

Appendix Y. Summary of statistical comparisons of dominant families abundances using Welch's t-test with Bonferroni correction to evaluate differences between native grassland and cultivated plots.

Family	Cultivated plots		Native grassland plot		p-value	Corrected p-value	Difference between means
	Mean rel. freq. (%)	Std. dev. (%)	Mean rel. freq. (%)	Std. dev. (%)			
Bacillaceae	8.038	1.651	0.357	0.470	0.000	0.000	7.681
Planococcaceae	2.676	0.615	0.029	0.034	0.000	0.000	2.646
Microbacteriaceae	2.417	0.623	0.018	0.021	0.000	0.000	2.399
Gemmatimonadaceae	1.472	0.374	0.206	0.082	0.000	0.000	1.266
Pirellulaceae	2.601	0.688	0.196	0.102	0.000	0.000	2.405
Uncultured Vicinamibacterales	2.995	0.652	0.831	0.348	0.000	0.000	2.164
Vicinamibacteraceae	2.106	0.635	0.000	0.000	0.000	0.000	2.106
Nocardiodaceae	1.569	0.495	0.000	0.000	0.000	0.000	1.569
67-14	1.677	0.576	0.000	0.000	0.000	0.000	1.677
Isosphaeraceae	0.905	0.548	3.478	0.495	0.000	0.000	-2.573
Micrococcaceae	2.682	0.946	0.001	0.003	0.000	0.000	2.681
Subgroup_2	0.038	0.043	6.131	0.602	0.000	0.000	-6.093
Chitinophagaceae	1.919	0.620	0.457	0.118	0.000	0.000	1.462
Gaiellaceae	1.692	0.710	0.001	0.002	0.000	0.000	1.691
Chthoniobacteraceae	1.055	0.360	4.419	0.702	0.000	0.000	-3.364
Solibacteraceae	0.297	0.310	2.553	0.509	0.000	0.000	-2.256
Gemmataceae	0.632	0.233	4.608	0.880	0.000	0.000	-3.975
Ktedonobacteraceae	0.117	0.165	5.846	1.386	0.000	0.001	-5.729
Xanthobacteraceae	3.213	1.118	14.996	3.629	0.000	0.005	-11.783
Xiphinematobacteraceae	0.030	0.066	4.607	1.383	0.000	0.007	-4.577
Beijerinckiaceae	0.945	0.294	4.972	1.257	0.000	0.007	-4.027
Uncultured Elsterales	0.186	0.179	8.280	2.520	0.000	0.008	-8.095
Solirubrobacteraceae	0.597	0.319	1.959	0.581	0.000	0.042	-1.362
Pseudomonadaceae	2.155	2.013	0.000	0.000	0.000	0.184	2.155
Pedosphaeraceae	1.143	0.338	0.788	0.266	0.010	4.944	0.355
Incertae Sedis	0.953	0.271	1.177	0.352	0.136	65.930	-0.224
Comamonadaceae	1.957	0.849	1.830	2.128	0.874	424.703	0.127

Appendix Z. Summary of statistical comparisons of dominant families abundances using Welch's t-test with Bonferroni correction to evaluate differences between the cultivated plots.

Family	3-year cultivated plot		30-year cultivated plot		p-value	Corrected p-value	Difference between means
	Mean rel. freq. (%)	Std. dev. (%)	Mean rel. freq. (%)	Std. dev. (%)			
Gemmatimonadaceae	1.813	0.178	1.131	0.125	0.000	0.000	0.683
Vicinamibacteraceae	1.596	0.453	2.617	0.282	0.000	0.052	-1.021
Chitinophagaceae	2.361	0.575	1.476	0.216	0.002	1.043	0.885
Solibacteraceae	0.522	0.292	0.072	0.074	0.002	1.064	0.450
Xanthobacteraceae	3.971	1.083	2.455	0.422	0.004	1.909	1.516
Isosphaeraceae	1.285	0.549	0.525	0.102	0.004	2.089	0.760
Uncultured Elsterales	0.306	0.178	0.065	0.055	0.005	2.276	0.242
Pirellulaceae	2.175	0.448	3.026	0.620	0.007	3.310	-0.852
Subgroup_2	0.065	0.041	0.012	0.024	0.008	3.793	0.053
Micrococcaceae	3.260	0.889	2.104	0.576	0.008	4.000	1.155
Gaiellaceae	1.262	0.437	2.122	0.668	0.009	4.283	-0.860
Chthoniobacteraceae	1.266	0.303	0.843	0.280	0.010	5.080	0.423
Ktedonobacteraceae	0.216	0.185	0.018	0.023	0.016	7.941	0.198
Solirubrobacteraceae	0.741	0.353	0.452	0.193	0.064	31.000	0.289
Beijerinckiaceae	1.069	0.281	0.822	0.251	0.082	40.092	0.247
Xiphinematobacteraceae	0.058	0.084	0.003	0.005	0.098	47.494	0.056
Nocardiodaceae	1.395	0.414	1.743	0.508	0.153	74.370	-0.348
Unknown_Family	0.862	0.237	1.043	0.273	0.176	85.713	-0.181
Pedosphaeraceae	1.239	0.393	1.047	0.237	0.259	125.877	0.191
Microbacteriaceae	2.265	0.542	2.568	0.661	0.331	160.771	-0.303
Gemmataceae	0.579	0.252	0.685	0.198	0.364	176.914	-0.106
Pseudomonadaceae	1.760	0.835	2.550	2.664	0.443	215.198	-0.790
Bacillaceae	7.807	1.650	8.268	1.619	0.581	282.188	-0.461
Uncultured Vicinamibacterales	2.939	0.660	3.051	0.639	0.733	356.016	-0.113
Planococcaceae	2.625	0.338	2.727	0.798	0.744	361.823	-0.102
67-14	1.633	0.648	1.721	0.490	0.764	371.153	-0.088
Comamonadaceae	2.019	0.865	1.895	0.827	0.772	375.372	0.125

Appendix AA. Summary of statistical comparisons of dominant genera abundances using Welch's t-test with Bonferroni correction to evaluate differences between native grassland and cultivated plots.

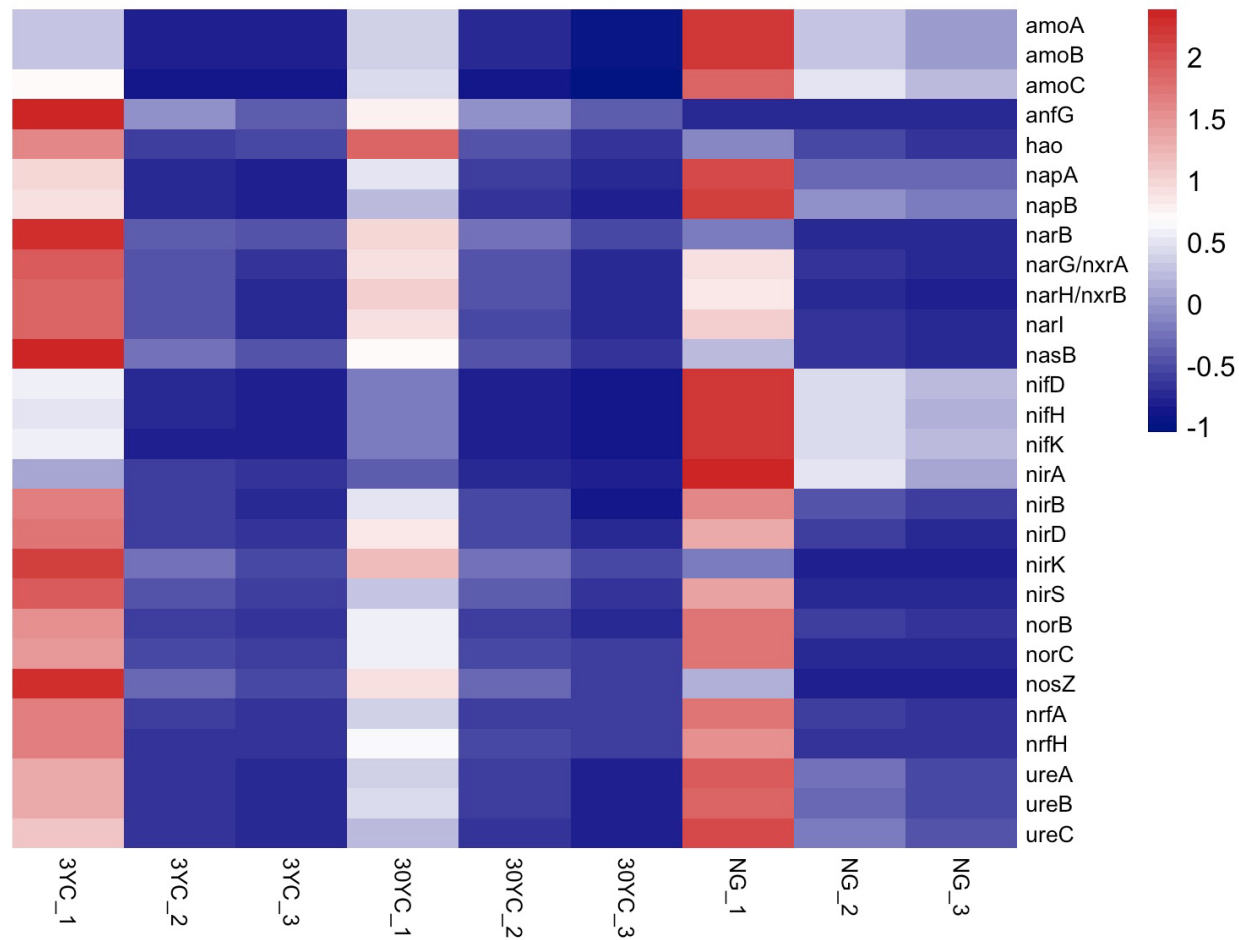
Genus	Cultivated plots		Native grassland plot		p-value	Corrected p-values	Difference between means
	Mean rel. freq. (%)	Std. dev. (%)	Mean rel. freq. (%)	Std. dev. (%)			
Unassigned Bacillaceae	4.020	1.105	0.000	0.000	0.000	0.000	4.020
Unassigned Vicinamibacteraceae	1.739	0.491	0.000	0.000	0.000	0.000	1.739
Uncultured Vicinamibacterales	2.995	0.652	0.831	0.348	0.000	0.000	2.164
Unassigned Microbacteriaceae	1.650	0.495	0.005	0.016	0.000	0.000	1.645
67-14	1.677	0.576	0.000	0.000	0.000	0.000	1.677
Subgroup_2	0.038	0.043	6.131	0.602	0.000	0.000	-6.093
Gaiella	1.692	0.710	0.001	0.002	0.000	0.000	1.691
Candidatus Udaeobacter	0.594	0.245	4.323	0.714	0.000	0.000	-3.728
Candidatus Solibacter	0.297	0.310	2.553	0.509	0.000	0.000	-2.256
Roseiarcus	0.032	0.061	3.685	0.706	0.000	0.000	-3.653
Uncultured Gemmataceae	0.290	0.176	4.545	0.870	0.000	0.001	-4.255
HSB_OF53-F07	0.018	0.041	3.728	0.872	0.000	0.002	-3.710
Uncultured Xanthobacteraceae	0.986	0.689	9.633	2.224	0.000	0.002	-8.647
Candidatus Xiphinematobacter	0.030	0.066	4.607	1.383	0.000	0.013	-4.577
Bradyrhizobium	0.872	0.346	2.868	0.684	0.000	0.014	-1.996
Uncultured Elsterales	0.186	0.179	8.280	2.520	0.000	0.016	-8.095
Pseudomonas	2.155	2.013	0.000	0.000	0.000	0.353	2.155
Bacillus	2.053	1.681	0.357	0.470	0.001	0.812	1.696
Unassigned Xanthobacteraceae	0.940	0.332	1.984	0.840	0.008	7.169	-1.044

Appendix AB. Summary of statistical comparisons of dominant genera abundances using Welch's t-test with Bonferroni correction to evaluate differences between the cultivated plots.

Genus	3-year cultivated plot		30-year cultivated plot		p-value	Corrected p-value	Difference between means
	Mean rel. freq. (%)	Std. dev. (%)	Mean rel. freq. (%)	Std. dev. (%)			
Unassigned Vicinamibacteraceae	1.379	0.374	2.099	0.288	0.001	0.570	-0.721
Unassigned Bacillaceae	3.245	0.715	4.795	0.854	0.001	1.161	-1.550
Candidatus Solibacter	0.522	0.292	0.072	0.074	0.002	2.038	0.450
Uncultured Xanthobacteraceae	1.474	0.650	0.498	0.224	0.003	2.351	0.976
Uncultured Elsterales	0.306	0.178	0.065	0.055	0.005	4.360	0.242
Candidatus Udaeobacter	0.749	0.226	0.439	0.143	0.006	5.362	0.310
Subgroup_2	0.065	0.041	0.012	0.024	0.008	7.265	0.053
Gaiella	1.262	0.437	2.122	0.668	0.009	8.206	-0.860
Bacillus	3.140	1.793	0.966	0.270	0.009	8.250	2.174
Bradyrhizobium	1.064	0.305	0.680	0.270	0.017	15.931	0.384
Roseiarcus	0.064	0.073	0.000	0.000	0.039	36.161	0.064
Candidatus Xiphinematobacter	0.058	0.084	0.003	0.005	0.098	90.982	0.056
HSB_OF53-F07	0.028	0.053	0.008	0.017	0.342	318.145	0.020
Pseudomonas	1.760	0.835	2.550	2.664	0.443	412.241	-0.790
Unassigned Xanthobacteraceae	0.893	0.414	0.986	0.212	0.582	541.664	-0.093
Uncultured Gemmataceae	0.269	0.182	0.311	0.167	0.636	592.338	-0.042
Unassigned Microbacteriaceae	1.705	0.452	1.596	0.529	0.664	617.917	0.109
Uncultured Vicinamibacterales	2.939	0.660	3.051	0.639	0.733	681.997	-0.113
67-14	1.633	0.648	1.721	0.490	0.764	710.995	-0.088

Appendix AC. Scaled abundance of nitrogen cycle genes across sampling plots and months.

3YC: 3-year cultivated plot. 30YC: 30-year cultivated plot. NG: Native grassland plot. Numbers next to the sampling plot labels indicate the sampling month.



Appendix AD. Summary of statistical comparisons of pathways abundance between native grassland and cultivated plots.

Pathway	Statistical comparisons		Shapiro Wilk test p-value		Levenne test p-value	Significance
	Test	p-value	Native grassland plot	Cultivated plots		
Assimilatory nitrate to nitrite reduction	T-test	0.000	0.455	0.673	0.064	Significant
Dissimilatory nitrate reduction or nitrite oxidation	T-test	0.000	0.907	0.231	0.060	Significant
Dissimilatory nitrite to ammonium reduction	T-test	0.000	0.991	0.447	0.403	Significant
Ammonification	Wilcoxon test	0.000	0.018	0.170	0.493	Significant
Assimilatory nitrite to ammonium reduction	Wilcoxon test	0.000	0.028	0.067	0.038	Significant
Nitrite to nitric oxide reduction	Wilcoxon test	0.000	0.003	0.971	0.798	Significant
Nitrogen fixation	Wilcoxon test	0.000	0.089	0.096	0.003	Significant
Nitrous oxide to nitrogen reduction	Wilcoxon test	0.000	0.008	0.685	0.618	Significant
Ammonia oxidation	Wilcoxon test	0.001	0.060	0.966	0.016	Significant
Nitric oxide to nitrous oxide reduction	Wilcoxon test	0.145	0.001	0.079	0.349	Not significant
Dissimilatory nitrate to nitrite reduction	T-test	0.210	0.383	0.619	0.418	Not significant

Appendix AE. Detailed results of the two-way repeated-measures ANOVA and Tukey HSD *post-hoc* comparisons for the absolute abundances of N-cycling genes

Gene: *nifH*

Error: Point

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Plot	2	1.738	0.8690	26.6	0.00104 **
Residuals	6	0.196	0.0327		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: Point:Month

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Month	2	0.4097	0.20486	7.831	0.00666 **
Plot:Month	4	0.1597	0.03992	1.526	0.25631
Residuals	12	0.3139	0.02616		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

contrast	estimate	SE	df	t.ratio	p.value
30-year cult. vs 3-year cult.	-0.126	0.0852	6	-1.480	0.3638
30-year cult. vs native grass.	-0.590	0.0852	6	-6.925	0.0011
3-year cult. vs native grass.	-0.464	0.0852	6	-5.445	0.0038

Gene: *nosZ*

Error: Point

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Plot	2	1.3269	0.6635	14.29	0.00523 **
Residuals	6	0.2786	0.0464		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: Point:Month

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Month	2	0.05857	0.029285	1.298	0.309
Plot:Month	4	0.01806	0.004515	0.200	0.933
Residuals	12	0.27075	0.022563		

contrast	estimate	SE	df	t.ratio	p.value
30-year cult. vs 3-year cult.	-0.144	0.102	6	-1.414	0.3926
30-year cult. vs native grass.	-0.525	0.102	6	-5.171	0.0050
3-year cult. vs native grass.	-0.382	0.102	6	-3.758	0.0220

=====

Gene: *nirS*

=====

Error: Point

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Plot	2	1.756	0.8779	10.53	0.0109 *
Residuals	6	0.500	0.0833		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: Point:Month

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Month	2	0.1869	0.09346	0.738	0.499
Plot:Month	4	0.2510	0.06276	0.496	0.739
Residuals	12	1.5193	0.12661		

contrast	estimate	SE	df	t.ratio	p.value
30-year cult. vs 3-year cult.	0.0317	0.136	6	0.233	0.9707
30-year cult. vs native grass.	0.5561	0.136	6	4.086	0.0152
3-year cult. vs native grass.	0.5244	0.136	6	3.854	0.0197

=====

Gene: *amoA* (AOA)

=====

Error: Point

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Plot	2	2.727	1.3634	2.721	0.144
Residuals	6	3.007	0.5011		

Error: Point:Month

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Month	2	1.7951	0.8976	5.020	0.0261 *
Plot:Month	4	0.9952	0.2488	1.391	0.2949
Residuals	12	2.1456	0.1788		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

=====

Gene: *amoA* (AOB)

=====

Error: Point

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Plot	2	128.24	64.12 5	2.45	0.000158 ***
Residuals	6	7.34	1.22		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: Point:Month

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Month	2	4.464	2.232	2.267	0.1462
Plot:Month	4	14.844	3.711	3.768	0.0329 *
Residuals	12	11.817	0.985		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

contrast	estimate	SE	df	t.ratio	p.value
30-year cult. vs 3-year cult.	0.0498	0.521	6	0.096	0.9950
30-year cult. vs native grass.	4.6478	0.521	6	8.917	0.0003
3-year cult. vs native grass.	4.5980	0.521	6	8.821	0.0003

Plot = 30-year cult.

contrast	estimate	SE	df	t.ratio	p.value
Month1 - Month2	0.6176	0.81	12	0.762	0.7323
Month1 - Month3	0.1122	0.81	12	0.139	0.9895
Month2 - Month3	-0.5053	0.81	12	-0.624	0.8102

Plot = 3-year cult.

contrast	estimate	SE	df	t.ratio	p.value
Month1 - Month2	0.8682	0.81	12	1.071	0.5486
Month1 - Month3	-0.1519	0.81	12	-0.188	0.9808
Month2 - Month3	-1.0201	0.81	12	-1.259	0.4434

Plot = native grass.

contrast	estimate	SE	df	t.ratio	p.value
Month1 - Month2	-2.8566	0.81	12	-3.526	0.0108
Month1 - Month3	-2.9451	0.81	12	-3.635	0.0089
Month2 - Month3	-0.0885	0.81	12	-0.109	0.9934

Appendix AF. Spearman correlation coefficient the gene abundance predicted by PICRUSt2 and the gene abundance quantified by qPCR

	predicted <i>nifH</i>	predicted <i>nirS</i>	predicted <i>nosZ</i>	predicted <i>amoA</i>
quantified <i>nifH</i>	0.811 (p < 0.001)			
quantified <i>nirS</i>		0.363 (p = 0.06)		
quantified <i>nosZ</i>			-0.239 (p = 0.23)	
quantified <i>amoA</i> AOA				0.055 (p = 0.78)
quantified <i>amoA</i> AOB				-0.437 (p = 0.03)

References

- Fisher, R., Hobgen, S., Mandaya, I., Kaho, N.R., Zulkarnain, 2017. Satellite image analysis and terrain modelling. A practical manual for natural resource management, disaster risk and development planning using free geospatial data and software. Charles Darwin University, Universitas Nusa Cendana dan Universitas Halu Oleo, Darwin, Australia.
- IGAC, 2015. Levantamiento Semidetallado de Suelos en las Áreas de Influencia de los Páramos de Colombia. Distrito de Páramos Santander. Instituto Geográfico Agustín Codazzi, Bogotá, Colombia.
- Laverman, A., Speksnijder, A., Braster, M., Kowalchuk, G., Verhoef, H., Van Verseveld, H., 2001. Spatiotemporal stability of an ammonia-oxidizing community in a nitrogen-saturated forest soil. *Microbial Ecology* 42, 35–45.
- Levy-Booth, D.J., Prescott, C.E., Grayston, S.J., 2014. Microbial functional genes involved in nitrogen fixation, nitrification and denitrification in forest ecosystems. *Soil Biology and Biochemistry* 75, 11–25. <https://doi.org/10.1016/j.soilbio.2014.03.021>
- Liu, M., Xue, R., Han, N., Yang, S., Wang, D., Hu, Y., Gu, K., Su, J., 2024. The impact of different preceding crops on soil nitrogen structure and nitrogen cycling in tobacco-planting soil. *Scientific Reports* 14. <https://doi.org/10.1038/s41598-024-52285-z>
- Moore, I.D., Grayson, R.B., Ladson, A.R., 1991. Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes* 5, 3–30. <https://doi.org/10.1002/hyp.3360050103>
- Rodríguez, G., Correa, A.M., Arango, M.I., Zapata, G., 2020. Catálogo de unidades litoestratigráficas de Colombia: Monzogranito de Santa Bárbara, Cordillera Oriental, Macizo de Santander, departamento de Santander. Servicio Geológico Colombiano.
- Rösch, C., Mergel, A., Bothe, H., 2002. Biodiversity of denitrifying and dinitrogen-fixing bacteria in an acid forest soil. *Applied and Environmental Microbiology* 68, 3818–3829. <https://doi.org/10.1128/AEM.68.8.3818-3829.2002>
- Rothauwe, J.-H., Witzel, K.-P., Liesack, W., 1997. The ammonia monooxygenase structural gene amoA as a functional marker: Molecular fine-scale analysis of natural ammonia-oxidizing populations. *Applied and Environmental Microbiology* 63, 4704–4712. <https://doi.org/10.1128/aem.63.12.4704-4712.1997>
- Throbäck, I.N., Enwall, K., Jarvis, A., Hallin, S., 2004. Reassessing PCR primers targeting nirS, nirK and nosZ genes for community surveys of denitrifying bacteria with DGGE. *FEMS Microbiology Ecology* 49, 401–417. <https://doi.org/10.1016/j.femsec.2004.04.011>
- Zevenbergen, L.W., Thorne, C.R., 1987. Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms* 12, 47–56. <https://doi.org/10.1002/esp.3290120107>