

Supplement of:

Near-tree soil footprints show higher organic carbon and lower bulk density with no detectable end-of-summer moisture penalty

Ana Paula Ferreira de Brito^{1,*}, Mark Farrell², Luke Mosley¹

1) School of Agriculture, Food and Wine, Adelaide University, Waite Campus, Glen Osmond, SA, Australia

2) CSIRO Agriculture & Food, Gate 4, 40 Waite Road, Urrbrae SA 5064, Australia.

*Corresponding author: Email anapaula.ferreiradebrito@adelaide.edu.au

Supplementary Material

Supplementary site and sampling information

Table S1. Australian Soil Classification (ASC) group and subgroup designations for the four cohorts, based on field-profile observations at the Waite Arboretum study site (Isbell and the National Committee on Soil and Terrain, 2021). Field descriptions included texture, colour, structure and diagnostic horizons; laboratory confirmation remains pending for selected chemical and structural properties.

Cohorts	Key Morphological Features	ASC Classification	Subgroup	Comments
C1-C4	Clear clay increases at ~25 cm; abrupt texture contrast between sandy loam A horizon and clayey B horizon; redder subsoil with moderate blocky structure and roots present	Red Chromosol	Mesotrophic / Eutrophic	Texture-contrast profile with a clay-rich B horizon and moderate structure, consistent with site soils observed at the study site

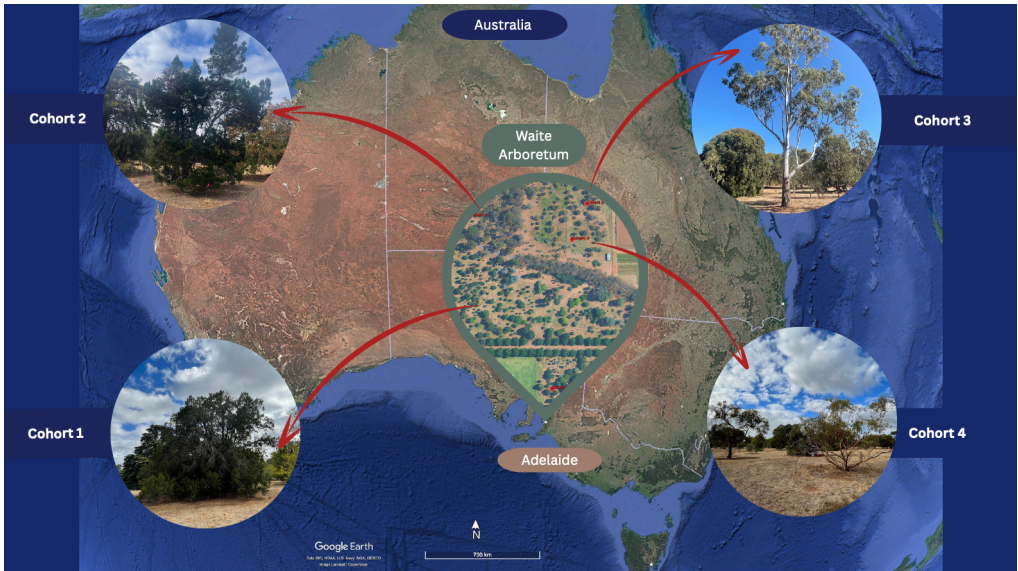


Figure S1. Study site and cohort overview at the Waite Arboretum (Adelaide, South Australia). The central map shows the location of the Waite Arboretum within Australia and the study area within the arboretum (pin/inset). Surrounding photographs provide representative examples of the four tree cohorts (C1–C4) sampled in this study, illustrating the contrasting tree types and open inter-canopy cropland context used in the depth × distance sampling design. Imagery © 2025 Landsat / Copernicus, Data SIO, NOAA, U.S. Navy, NGA, GEBCO; Map data © 2025 Google.

Table S2. Summary of the sampling design and experimental setup at the Waite Arboretum study site. Mineral-soil depth grids were response-specific: bulk density at the two directly measured 0–5 and 5–10 cm ring-core layers; gravimetric moisture at four layers (0–5, 5–10, 10–20 and 20–30 cm); and SOC, total N, C:N, mineral N and Colwell-extractable P in auger samples at three layers (0–10, 10–20 and 20–30 cm), which also served as the fixed-depth SOC-stock layers

Parameter	Details
Study site	Waite Arboretum, Adelaide, South Australia; Mediterranean climate; 618 mm mean annual rainfall and 21.4 °C mean annual maximum temperature as supplied in the study metadata.
Cohorts	Four tree cohorts (C1-C4) representing different tree groups occurring in dry South Australian landscapes.
Trees per cohort	Three to four individual trees per cohort; 14 trees in total.
Tree attributes recorded	Species, family, age, height, canopy area, total stored C (kg) and annual C sequestration (kg yr ⁻¹).
Transects	Four transect directions per tree: NE, SE, SW and NW.
Sampling distances	1, 2, 4 and 8 m from the tree stem along each transect.
Response-specific depth grids	Litter at the surface. Mineral-soil grids differed by response: bulk density at 0–5 and 5–10 cm (2 layers; 14 × 4 × 2 = 112 composites); gravimetric moisture at 0–5, 5–10, 10–20 and 20–30 cm (4 layers; 14 × 4 × 4 = 224 tree-level observations); and SOC, total N, C:N, mineral N and Colwell-extractable P in auger samples at 0–10, 10–20 and 20–30 cm (3 layers; 14 × 4 × 3 = 168 tree-level observations). Values were pooled from the four transect subsamples within each tree × distance × depth cell.
Fixed-depth SOC stock grid	Composite 0-10, 10-20 and 20-30 cm layers; 56 complete tree-distance profiles and 168 layer-specific stock observations.
Bulk-density cores	Two stainless-steel rings (5 cm height × 5 cm diameter) for each transect × distance combination at 0-5 and 5-10 cm, giving 448 rings. Rings were composited by Tree × Distance × Depth to yield 112 bulk-density samples.
Litter samples	One field sample at each Tree × Transect × Distance position (14 × 4 × 4 = 224), composited by Tree × Distance to yield 56 litter samples.
Auger samples for fixed-depth stocks	Four transect samples per Tree × Distance position were sectioned into 0-10, 10-20 and 20-30 cm intervals and pooled within Tree × Distance × Depth, producing 168 composite mineral soil samples.
Storage	Samples were labelled in the field and stored under refrigeration before laboratory processing.



Figure S2. Field sampling and soil sample handling workflow. (a) Example soil pit/auger hole showing sampling depth in the field. (b) Bulk density core rings and associated bagged soil samples labelled for laboratory processing. (c) Representative soil samples laid out for documentation and visual assessment of colour and texture before analysis.

Supplementary laboratory details

Bulk density (ρ_b) and coarse fragment corrections

Undisturbed soil cores ($n = 112$ tree-distance-depth composites) were dried at 105 °C to constant mass. Core ring volume (V_{core} , cm³) was calculated geometrically:

$$V_{core} = \pi r^2 h$$

Whole-soil bulk density, including the > 2mm fraction, was calculated as:

$$\rho_{b_{total}} = \frac{M_{total}}{V_{core}}$$

where: M_{total} is the total oven-dry core mass (fine earth plus coarse fragments) and V_{core} is the total ring volume.

The volumetric coarse fragment fraction (CFv ; v/v) was calculated by converting coarse fragment mass (g) to volume (cm^3) using a particle density of 2.9 g cm^{-3} for the lateritic/ironstone fragments and dividing by total core volume (Harbo et al., 2022):

$$CFv = \frac{\left(\frac{M_{CF}}{\rho_{CF}}\right)}{V_{core}} \quad \text{with} \quad \rho_{CF} = 2.9 \text{ g m}^{-3}$$

Coarse fragments were corrected once in SOC stock calculations through the term $(1 - CFv)$, used together with whole-soil bulk density. Bulk density was not additionally converted to a fine-earth bulk density, thereby avoiding double correction.

Litter Carbon

Mean litter C concentration differed among cohorts on the primary raw scale ($F(3, 10) = 5.41$, $p = 0.018$; Supplementary Table S3a). Mean litter C ($\pm$ SE among trees) was highest in C3 ($424.6 \pm 19.7 \text{ g kg}^{-1}$), followed by C4 ($336.0 \pm 17.0 \text{ g kg}^{-1}$), C2 ($315.7 \pm 20.2 \text{ g kg}^{-1}$), and C1 ($295.8 \pm 30.2 \text{ g kg}^{-1}$). The parametric Holm-adjusted comparison indicated higher litter C in C3 than in C1 (difference = 128.8 g kg^{-1} ; adjusted $p = 0.019$), but this pairwise localization was not supported by the exact permutation sensitivity after correction across the six cohort comparisons (adjusted $p = 0.286$). The cohort effect was also weaker on the log-transformed scale ($p = 0.066$), indicating that the magnitude and pairwise localization of the cohort difference should be interpreted cautiously (Supplementary Table S4b). Litter C did not vary significantly with distance from the tree stem ($F(3, 30) = 0.78$, $p = 0.515$), and the primary categorical Distance $\times$ Cohort interaction was not significant ($F(9, 30) = 1.41$, $p = 0.227$). Both conclusions remained unchanged after sphericity correction and parametric-bootstrap

validation (Supplementary Table S4a). None of the three planned contrasts against 8 m remained significant after Holm adjustment (parametric adjusted $p \geq 0.376$; exact adjusted $p \geq 0.400$), and the overall continuous gradient was -0.44 g kg^{-1} per doubling of distance (95% CI: -14.81 to 13.93 ; $t(10) = -0.07$, $p = 0.947$). A secondary continuous-trend analysis nevertheless indicated heterogeneity among cohort-specific slopes ($F(3, 10) = 5.37$, $p = 0.018$); because the primary categorical interaction was not significant, this result was treated as exploratory (Supplementary Table S4b). Descriptively, mean litter C was 5.8, 9.8, 17.1 and 12.7 times the corresponding 0–10 cm mineral soil SOC concentration in C1, C2, C3 and C4, respectively.

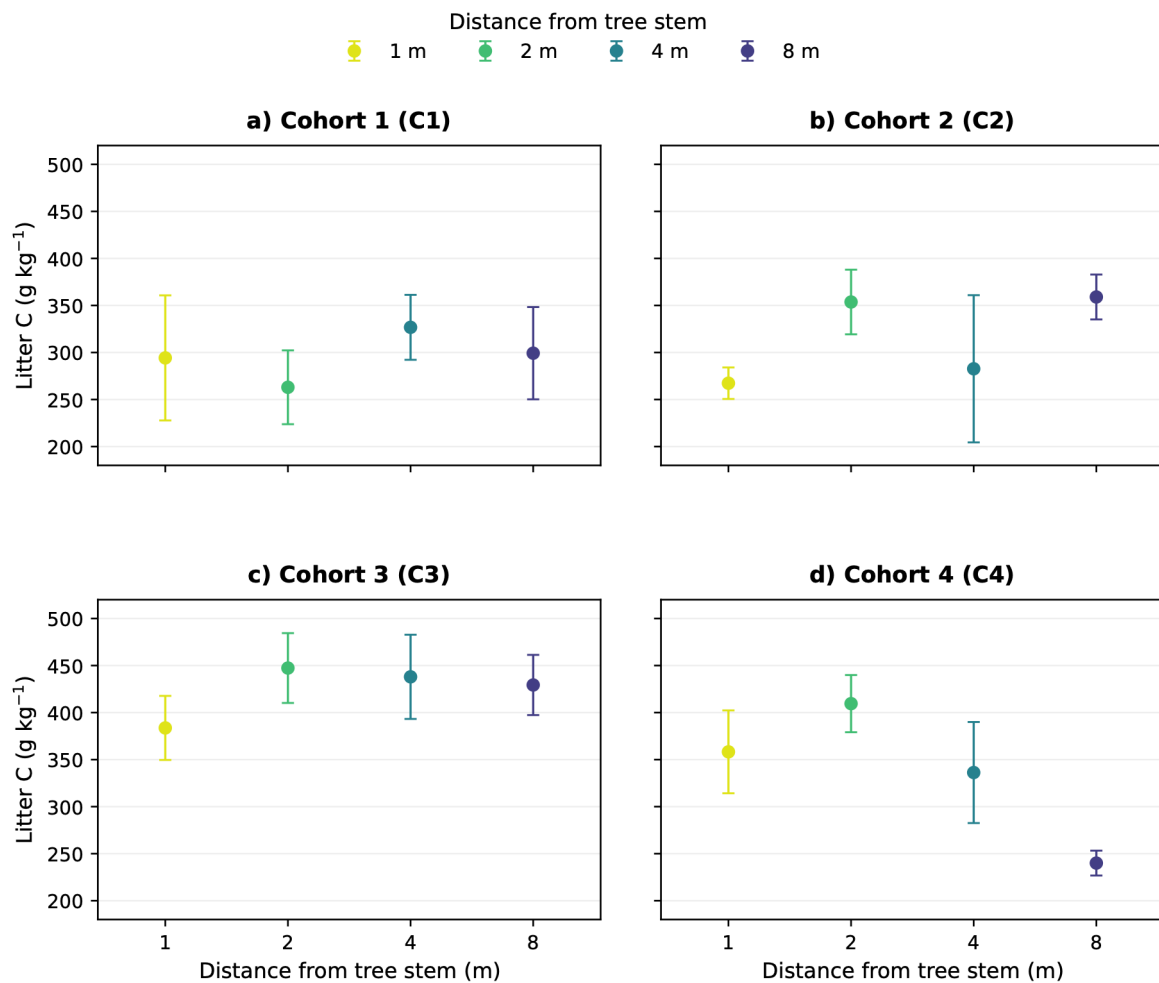


Figure S3. Litter carbon concentration across distance and cohort. Points show arithmetic means and error bars represent ± 1 SE among independent trees within each cohort and distance combination (C1 and C4: $n = 4$ trees; C2 and C3: $n = 3$ trees). Colours represent distance from the tree stem. The figure is descriptive; statistical inference was based on tree-level split-plot tests and Holm-adjusted planned contrasts.

Total Nitrogen

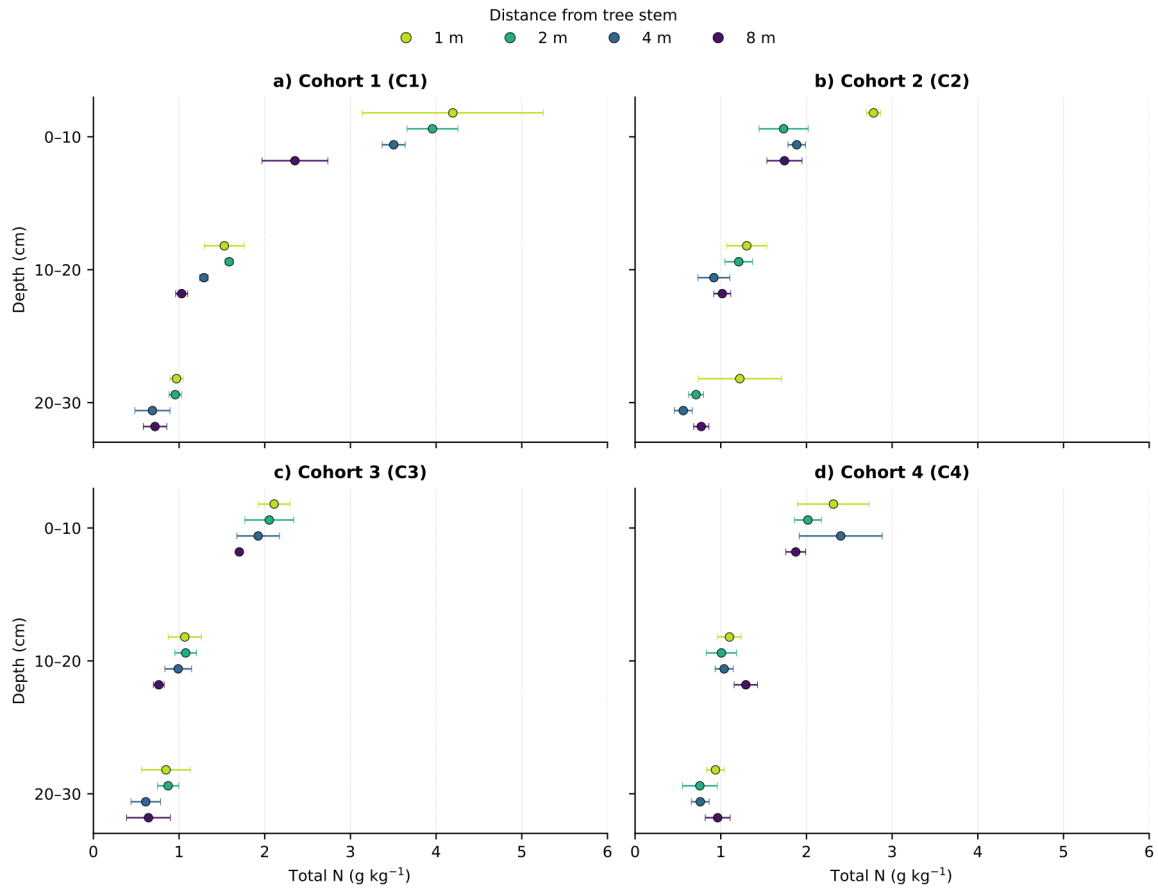


Figure S4. Total nitrogen (TN) concentrations across the auger-sampled 0–10, 10–20 and 20–30 cm layers and distances from the tree stem for each cohort. Points show observed means and error bars represent SE among trees (C1 and C4: n = 4; C2 and C3: n = 3); colours indicate distances of 1, 2, 4 and 8 m from the tree stem.

Gravimetric and hygroscopic moisture

Fresh soil subsamples were weighed, oven-dried at 105 °C to constant mass, and reweighed.

Gravimetric moisture was computed as:

$$\theta_m = \frac{(M_{\text{moist}} - M_{\text{dry}})}{M_{\text{dry}}}$$

and expressed as percentage by multiplying by 100.

SOC and total N concentrations were determined and analysed on material dried below 40° C.

The hygroscopic correction was applied only to place SOC concentrations on the 105° C oven-

dry basis required for stock calculations; it was not applied to the SOC or total N concentration analyses. For a representative subset, hygroscopic moisture was calculated as:

$$\theta_h = \frac{M_{40} - M_{105}}{M_{105}}$$

The depth-specific factors used for stocks were 0.0120 for 0-10 cm, 0.0100 for 10-20 cm and 0.0085 for 20-30 cm. SOC concentration on an oven-dry basis was calculated as:

$$SOC_{OD} = SOC_{40} \times (1 + \theta_h)$$

Fresh-soil gravimetric moisture was analysed as a separate ecological response and was not used as θ_h in the SOC stock equation.

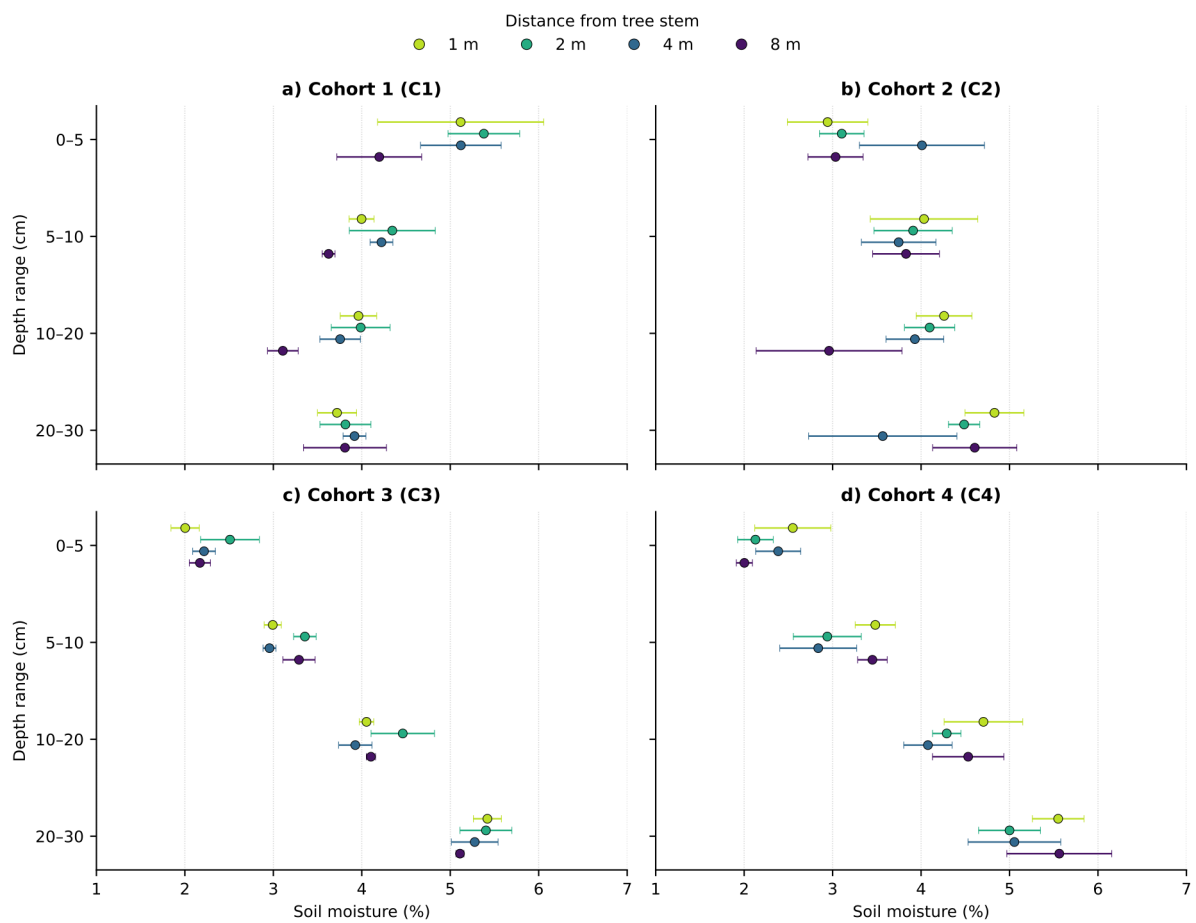


Figure S5. Gravimetric soil moisture across four mineral soil depth layers, four distances from the tree stem and the four cohorts. Points show observed tree-level means and error bars show ± 1 SE among independent trees; colours represent distances of 1, 2, 4 and 8 m.

Supplementary Methods S1. Statistical analysis of soil properties

S1.1 Experimental unit and split-plot structure

All final calculations were implemented in Python 3.13.9 from the local Anaconda distribution. Reproducible Python scripts and matching Jupyter Notebook 7.3.2 notebooks were retained, and all tables and figures were generated from the same analysis functions. The verified software environment included Anaconda Navigator 2.7.0 and statsmodels 0.14.4 where applicable (Seabold and Perktold, 2010). For mineral-soil responses, Cohort (C1–C4) was treated as a between-tree factor, whereas Distance from the tree stem (1, 2, 4 and 8 m) and the response-specific Depth factor were repeated within tree. Primary BD inference used the directly measured 0–5 and 5–10 cm ring cores; gravimetric moisture used 0–5, 5–10, 10–20 and 20–30 cm; and SOC, total N, C:N, mineral-N and Colwell-P concentration analyses used auger samples at 0–10, 10–20 and 20–30 cm. Tree identity, defined by the unique Cohort × Tree combination, was the experimental and repeated-sampling unit (14 trees; C1 = 4, C2 = 3, C3 = 3 and C4 = 4 trees). Litter C was analysed separately using the analogous structure, with Cohort as the between-tree factor and Distance as the repeated within-tree factor. Omnibus inference used tree-level split-plot F-tests with the corresponding error strata. Cohort was tested against variation among trees within cohorts. Distance, Depth and Distance × Depth, together with their interactions with Cohort, were tested against the corresponding within-tree error strata. The mineral soil models therefore evaluated Cohort, Distance, Distance × Cohort, Depth, Depth × Cohort, Distance × Depth and Distance × Depth × Cohort. Repeated observations from the same tree were not treated as independent replicates.

S1.2 Response scales, planned contrasts and multiplicity control

Response scales were selected from residual diagnostics. Directly measured bulk density, NO_2^- -N and litter C were analysed on their original scales. SOC, total N and C:N ratio were natural-log transformed; gravimetric soil moisture was square-root transformed; and NH_4^+ -N, NO_3^- -N and Colwell-extractable P were transformed as $\log_{10}(x)$. The $\log_{10}$ transformation retained the single observed zero NO_3^- -N value without introducing an arbitrary replacement. Untransformed tree-level means and standard errors are presented in figures and

descriptive summaries. Depth contrasts and any exploratory cohort-specific contrasts were adjusted within the prespecified family identified in their corresponding supplementary table. Holm adjustment (Holm, 1979) was used because the hypotheses comprised planned reference comparisons or explicitly specified pairwise families rather than all possible post-hoc comparisons. Statistical significance was assessed at $\alpha = 0.05$.

S1.3 Primary univariate results

Table S3 reports the corrected primary split-plot tests on each response-specific grid. Distance was significant for directly measured surface BD, SOC, total N, NO_3^- -N and Colwell-extractable P, but not for C:N, soil moisture, NH_4^+ -N, NO_2^- -N or litter C. No Distance $\times$ Cohort interaction was significant. Depth was significant for every mineral-soil response. Vertical stratification differed among cohorts for SOC, soil moisture, NH_4^+ -N, NO_2^- -N and Colwell P. The significant Distance $\times$ Depth and Distance $\times$ Depth $\times$ Cohort terms for Colwell P indicate that its radial pattern depended on depth and, jointly, cohort; these interactions should be interpreted hierarchically.

Table S3a. Primary tree-level split-plot tests for cohort and radial effects. Analysis scales are shown because inferential tests were conducted on the transformed scale where specified. Values are F (df₁, df₂) and p. Litter C has no mineral soil depth term.

Response	Analysis scale	Cohort	Distance	Distance $\times$ Cohort
Bulk density measured 0-10 cm	Original	F(3,10) = 6.37, p = 0.011	F(3,30) = 14.00, p < 0.001	F(9,30) = 0.79, p = 0.624
SOC concentration	ln(x)	F(3,10) = 3.65, p = 0.052	F(3,30) = 3.69, p = 0.023	F(9,30) = 1.45, p = 0.210
Total N	ln(x)	F(3,10) = 3.68, p = 0.051	F(3,30) = 3.67, p = 0.023	F(9,30) = 1.34, p = 0.257
C:N ratio	ln(x)	F(3,10) = 0.76, p = 0.540	F(3,30) = 0.10, p = 0.958	F(9,30) = 0.91, p = 0.529
Gravimetric soil moisture	sqrt(x)	F(3,10) = 1.83, p = 0.205	F(3,30) = 1.71, p = 0.185	F(9,30) = 1.37, p = 0.245
NH_4^+ -N	log1p(x)	F(3,10) = 13.71, p < 0.001	F(3,30) = 2.79, p = 0.058	F(9,30) = 1.47, p = 0.204
NO_2^- -N	Original	F(3,10) = 13.64, p < 0.001	F(3,30) = 0.45, p = 0.722	F(9,30) = 0.49, p = 0.869
NO_3^- -N	log1p(x)	F(3,10) = 3.33, p = 0.065	F(3,30) = 4.16, p = 0.014	F(9,30) = 0.98, p = 0.477
Colwell-extractable P	log1p(x)	F(3,10) = 17.84, p < 0.001	F(3,30) = 6.13, p = 0.002	F(9,30) = 1.82, p = 0.105
Litter C	Original	F(3,10) = 5.41, p = 0.018	F(3,30) = 0.78, p = 0.515	F(9,30) = 1.41, p = 0.227

Table S3b. Primary tree-level split-plot tests for depth and crossed within-tree effects. Values are F (df₁, df₂) and p.

Response	Depth	Depth $\times$ Cohort	Distance $\times$ Depth	Distance $\times$ Depth $\times$ Cohort
Bulk density 0-10 cm	F(1,10) = 208.27, p < 0.001	F(3,10) = 1.48, p = 0.279	F(3,30) = 1.17, p = 0.337	F(9,30) = 0.65, p = 0.750
SOC concentration	F(2,20) = 574.14, p < 0.001	F(6,20) = 11.69, p < 0.001	F(6,60) = 0.68, p = 0.667	F(18,60) = 0.63, p = 0.859
Total N	F(2,20) = 199.35, p < 0.001	F(6,20) = 2.19, p = 0.087	F(6,60) = 1.35, p = 0.249	F(18,60) = 0.40, p = 0.982
C:N ratio	F(2,20) = 14.57, p < 0.001	F(6,20) = 2.13, p = 0.095	F(6,60) = 1.54, p = 0.182	F(18,60) = 0.92, p = 0.564
Gravimetric soil moisture	F(3,30) = 48.58, p < 0.001	F(9,30) = 24.75, p < 0.001	F(9,90) = 1.56, p = 0.140	F(27,90) = 0.82, p = 0.718
NH_4^+ -N	F(2,20) = 76.27, p < 0.001	F(6,20) = 5.23, p = 0.002	F(6,60) = 1.34, p = 0.252	F(18,60) = 0.70, p = 0.802
NO_2^- -N	F(2,20) = 20.94, p < 0.001	F(6,20) = 8.11, p < 0.001	F(6,60) = 0.36, p = 0.902	F(18,60) = 0.33, p = 0.994
NO_3^- -N	F(2,20) = 122.16, p < 0.001	F(6,20) = 2.05, p = 0.106	F(6,60) = 0.98, p = 0.449	F(18,60) = 1.45, p = 0.140
Colwell-extractable P	F(2,20) = 85.61, p < 0.001	F(6,20) = 28.08, p < 0.001	F(6,60) = 2.80, p = 0.018	F(18,60) = 1.81, p = 0.045

S1.4 Sensitivity analyses for principal responses

The principal distance conclusions were evaluated using Greenhouse-Geisser correction (Greenhouse and Geisser, 1959), Huynh-Feldt-Lecoutre correction (Lecoutre, 1991) and 5,000-draw parametric-bootstrap

tests on each response-specific grid. Planned distance contrasts were additionally checked using exact tree-level sign-flip tests and continuous trends per doubling of distance. Leave-one-tree-out refitting was retained for BD, SOC, litter C and soil moisture. These analyses were treated as sensitivity checks rather than separate primary hypothesis tests.

Table S4a. Sphericity and parametric-bootstrap sensitivity analyses for the primary radial tests. All bootstrap tests used 5,000 draws.

Response	Term	Primary p	GG epsilon	GG p	HF-Lecoutre p	Bootstrap p
Bulk density (measured 0–10 cm)	Distance	<0.001	0.483	<0.001	<0.001	0.002
Bulk density (measured 0–10 cm)	Distance × Cohort	0.624	0.483	0.557	0.567	0.546
SOC concentration	Distance	0.023	0.849	0.030	0.023	0.029
SOC concentration	Distance × Cohort	0.210	0.849	0.225	0.210	0.220
Total N	Distance	0.023	0.599	0.050	0.039	0.050
Total N	Distance × Cohort	0.257	0.599	0.291	0.280	0.277
C:N ratio	Distance	0.958	0.615	0.889	0.922	0.924
C:N ratio	Distance × Cohort	0.529	0.615	0.503	0.514	0.493
NH ₄ ⁺ -N	Distance	0.058	0.599	0.093	0.080	0.090
NH ₄ ⁺ -N	Distance × Cohort	0.204	0.599	0.246	0.232	0.235
NO ₂ ⁻ -N	Distance	0.722	0.513	0.598	0.623	0.609
NO ₂ ⁻ -N	Distance × Cohort	0.869	0.513	0.766	0.788	0.774
NO ₃ ⁻ -N	Distance	0.014	0.622	0.034	0.025	0.036
NO ₃ ⁻ -N	Distance × Cohort	0.477	0.622	0.463	0.469	0.452
Colwell-extractable P	Distance	0.002	0.646	0.009	0.005	0.011
Colwell-extractable P	Distance × Cohort	0.105	0.646	0.149	0.128	0.140
Litter C	Distance	0.515	0.520	0.446	0.461	0.438
Litter C	Distance × Cohort	0.227	0.520	0.274	0.265	0.260
Gravimetric soil moisture	Distance	0.185	0.925	0.190	0.185	0.194
Gravimetric soil moisture	Distance × Cohort	0.245	0.925	0.251	0.245	0.245

Table S4b. Sensitivity and continuous-trend analyses for litter C concentration. Parametric p-values for planned contrasts were Holm-adjusted within their stated families; exact label-permutation p-values for cohort comparisons were adjusted across all six pairwise comparisons.

Analysis	Primary result	Sensitivity or secondary result	Interpretation
Cohort omnibus	Raw-scale $F(3,10) = 5.41$, $p = 0.018$	Parametric bootstrap $p = 0.020$; log-scale $p = 0.066$; leave-one-tree-out $p < 0.05$ in 12 of 14 runs	Primary raw-scale evidence, but sensitive to response scale and individual-tree omission.
C3 versus C1	Difference = 128.8 g kg^{-1} ; parametric Holm $p = 0.019$	Exact unadjusted $p = 0.057$; exact Holm $p = 0.286$	Pairwise localization was not confirmed by the exact permutation sensitivity.
Distance omnibus	$F(3,30) = 0.78$, $p = 0.515$	GG $p = 0.446$; HF-Lecoutre $p = 0.461$; bootstrap $p = 0.438$	No overall categorical distance effect.
Distance × Cohort	$F(9,30) = 1.41$, $p = 0.227$	GG $p = 0.274$; HF-Lecoutre $p = 0.265$; bootstrap $p = 0.260$	No primary categorical interaction.
Planned contrasts against 8 m	Minimum parametric Holm $p = 0.376$	Minimum exact sign-flip Holm $p = 0.400$	No planned distance contrast was supported.
Overall continuous slope	-0.44 g kg^{-1} per doubling; 95% CI: -14.81 to 13.93 ; $p = 0.947$	Exact sign-flip $p = 0.942$	No overall continuous distance trend.
Continuous slope × Cohort	$F(3,10) = 5.37$, $p = 0.018$	Mean tree-level slopes: C1 = $+7.9$, C2 = $+20.4$, C3 = $+12.8$ and C4 = -42.8 g kg^{-1} per doubling	Exploratory heterogeneity because the primary categorical interaction was not supported.

For TN, the corrected Distance test remained at the 0.05 boundary under GG correction and parametric bootstrap (both $p = 0.050$), while the HF-Lecoutre result was $p = 0.039$. For Colwell P, the parametric Holm-adjusted 1 m versus 8 m contrast was $p = 0.046$ and the exact sign-flip result was $p = 0.055$; for NO₃⁻-N the corresponding values were $p = 0.093$ and $p = 0.044$. These differences concern localization of the radial effects, not the omnibus conclusions, which were

supported by the sphericity-corrected and bootstrap tests in Table S4a. Complete contrast, continuous-gradient and response-scale results are retained in the response-specific reproducibility tables.

S1.5 Multivariate soil-property structure

Partial redundancy analysis (partial RDA) was performed entirely in Python using multivariate least squares. The response matrix comprised bulk density, gravimetric soil moisture, SOC, total N, NH_4^+ -N, NO_2^- -N, NO_3^- -N and Colwell-extractable P across the common grid of 14 trees $\times$ four distances $\times$ three auger-sampled depth layers ($n = 168$). At 0–10 cm, BD was the equal-thickness mean of the directly measured 0–5 and 5–10 cm ring cores; deeper BD was filled using the final model-assisted procedure. Each variable was analysed on the response scale used in the corresponding univariate analysis and then standardised to zero mean and unit sample standard deviation. Litter C, the separate 0–5 and 5–10 cm ring-core chemistry values, SOC and N stocks, and C:N ratios were excluded. Distance was tested after conditioning on Depth and Tree ID, and Depth was tested after conditioning on Distance and Tree ID. Interaction terms were evaluated hierarchically. Significance was assessed using 9,999 Freedman-Lane permutations (Freedman and Lane, 1983) restricted within Tree ID. Complete three-depth profiles were permuted among distances for distance-related tests, and complete four-distance profiles were permuted among depths for depth-related tests. A sensitivity matrix excluded model-assisted BD. This scheme retained the tree as the independent unit and preserved the crossed repeated-measures structure.

Table S5. Partial redundancy analysis of multivariate mineral soil structure. Unique R^2 is expressed relative to the total standardised multivariate variance. Permutation p-values were obtained using 9,999 Tree ID-restricted Freedman-Lane permutations.

Term	Conditioned on	df term	Residual df	Pseudo-F	Unique R^2	Unique adjusted R^2	Permutation p
Distance	TreeID + Depth	3	149	4.32	0.033	0.028	<0.001
Distance $\times$ Cohort	TreeID + Depth + Distance	9	140	1.45	0.033	0.011	0.173
Depth	TreeID + Distance	2	149	76.23	0.389	0.425	<0.001
Depth $\times$ Cohort	TreeID + Distance + Depth	6	143	8.25	0.098	0.096	<0.001
Distance $\times$ Depth	TreeID + Distance + Depth + Distance $\times$ Cohort + Depth $\times$ Cohort	6	128	0.99	0.011	-0.000	0.346
Distance $\times$ Depth $\times$ Cohort	TreeID + all lower-order terms	18	110	0.52	0.019	-0.022	0.950

Distance explained a small but statistically detectable component of multivariate soil-property variation after Depth and Tree ID were controlled. The Distance $\times$ Cohort interaction was not supported, whereas Depth and Depth $\times$ Cohort explained substantially larger components of multivariate variation. Excluding model-assisted BD retained the Distance effect (pseudo-F = 3.82, unique $R^2 = 0.032$, $p < 0.001$). Negative adjusted R^2 values for unsupported higher-order interactions reflect the penalty for model complexity and should not be interpreted as negative ecological variance.

Supplementary Methods S2. Depth-scaling pedotransfer calibration for bulk density and coarse fragments

S2.1 Purpose and data provenance

Whole-soil bulk density (ρ_b) and volumetric coarse-fragment fraction (CFv; v/v) were measured directly from intact cores at 0-5 and 5-10 cm and combined as equal-thickness means for the 0-10 cm stock layer. Direct intact-core measurements were unavailable for 10-20 and 20-30 cm. The 0-30 cm SOC stocks are therefore model-assisted below 10 cm.

The bulk-density reference data came from soil profiles sampled in the Waite Rotation Trial in 1997, before subsequent changes to rotation and management, and were independent of the arboretum trees sampled in the present study. The source table contained 68 subplot records at each depth (34 plots $\times$ eastern and western subplots). The 0-10, 10-20 and 20-40 cm bulk-density triplets were identical within every eastern-western subplot pair. Calibration, cross-validation and uncertainty estimation were consequently based on bulk-density profiles, rather than treating duplicated bulk-density values as independent observations. The 34 unique reference profiles are listed in Supplementary Table S9.

The reference dataset contained bulk density and volumetric carbon and nitrogen quantities but no texture or deep coarse-fragment measurements. Candidate bulk-density models were therefore restricted to non-circular predictors that were available in both the reference and arboretum datasets. Models using carbon concentration reconstructed as volumetric C divided by measured bulk density were examined diagnostically but were not eligible for selection because the reconstructed predictor was algebraically dependent on the response.

S2.2 Profile-wise model comparison and selection for 10-20 cm

Candidate models for the directly aligned 10-20 cm reference layer were compared using leave-one-profile-out cross-validation. The candidates included a training-layer mean, a training-layer median, the original 20% total trimmed depth ratio, linear and robust relationships with measured 0-10 cm bulk density, and a nested cross validated shrinkage of the ratio prediction towards the layer mean. Profile wise validation was used because systematic missingness by depth is more difficult than randomly missing individual horizons, and pedotransfer performance must be evaluated at the independent profile level (Arbor et al., 2023, 2024; Sequeira et al., 2014).

For every training fold, the trimmed ratio model was:

$$\rho_{b,target} = \rho_{b,0-10} \times \text{trimmean}(\rho_{b,target} / \rho_{b,0-10}),$$

where 20% was removed in total, equivalent to 10% from each tail. The full-data factor for 10-20 cm was 1.106531414.

Table S6a. Leave-one-profile-out performance for the aligned 10-20 cm layer and diagnostic performance of the former ratio at 20-40 cm. Bias is measured minus predicted bulk density.

Reference layer	Model	Profiles	RMSE (g cm ⁻³)	MAE (g cm ⁻³)	Median absolute error (g cm ⁻³)	Bias (g cm ⁻³)	Cross-validated R ²
10-20 cm	Selected: 20%-total trimmed ratio to measured 0-10 cm BD	34	0.110	0.081	0.054	-0.001	0.409
10-20 cm	Reference-layer mean benchmark	34	0.147	0.077	0.046	0.000	-0.062
20-40 cm	Reference-layer mean benchmark	34	0.075	0.049	0.039	0.000	-0.062
20-40 cm	Former: 20%-total trimmed ratio to measured 0-10 cm BD	34	0.238	0.100	0.043	-0.033	-9.645

The surface anchored ratio had the lowest cross validated (Root Mean Square Error) RMSE for 10-20 cm and was retained for that layer. At 20-40 cm, the former ratio model did not reproduce profile level variation, so the 20-40 observations were used as depth interval information for harmonisation rather than treated as direct 20-30 cm measurements.

S2.3 Mass preserving harmonisation from 20-40 to 20-30 cm

The reference dataset reported interval means for 0-10, 10-20 and 20-40 cm, whereas the arboretum stock calculation required 20-30 cm. The depth supports were harmonised

separately for every reference profile using a mass-preserving depth function (Bishop et al., 1999; Malone et al., 2009). For observed interval means y_i and boundaries z_i , cumulative depth integrals were constructed as:

$$I(z_0) = 0; I(z_j) = \sum y_i(z_i - z_{i-1}), \text{ for } i = 1, \dots, j.$$

A natural cubic spline was interpolated through the cumulative integrals at 0, 10, 20 and 40 cm. The derivative of this cumulative spline is a continuous quadratic depth function. Because the spline passes through every cumulative boundary value, the original interval means are preserved exactly. The harmonised 20-30 cm mean was:

$$\rho_{b, 20-30} = [I_{spline}(30) - I_{spline}(20)] / 10.$$

The maximum numerical discrepancy between an original reference-layer mean and its reconstructed mean was $6.7 \times 10^{-16} \text{ g cm}^{-3}$, and all predicted depth-function values remained positive. PCHIP (Piecewise Cubic Hermite Interpolating Polynomial) interpolation of the same cumulative integrals and the explicit assumption of constant bulk density within 20-40 cm were retained as model-form sensitivities.

An analogue validation was used because 20-30 cm was not directly observed in the reference dataset. The measured 10-20 and 20-40 values were first combined into a synthetic 10-40 cm mean; each method then used only 0-10 and synthetic 10-40 cm information to recover the directly measured 10-20 cm value. RMSE was the prespecified selection criterion, consistent with the candidate-model comparison.

Table S6b. Analog depth-disaggregation performance. The validation target was directly measured 10-20 cm bulk density after its lower boundary had been hidden within a synthetic 10-40 cm interval.

Harmonisation method	Profiles	RMSE (g cm ⁻³)	MAE (g cm ⁻³)	Bias (g cm ⁻³)	Cross-validated R ²	Maximum absolute error (g cm ⁻³)
Natural mass-preserving spline	34	0.109	0.073	0.061	0.423	0.445
Linear midpoint depth function	34	0.112	0.084	0.073	0.387	0.392
PCHIP mass-preserving interpolation	34	0.114	0.078	0.066	0.367	0.464
Constant within the composite interval	34	0.132	0.059	0.003	0.153	0.713

The natural mass-preserving spline had the lowest analog RMSE and was selected for the primary harmonisation. The constant-within-interval method had the lowest (Mean Absolute

Error) MAE and became the best RMSE method when the influential PeW_4 profile was omitted; consequently, method selection was not treated as evidence that one vertical curve was universally correct. The complete PCHIP, constant-within-20-40 and former smooth-curve calculations were retained as stock-level sensitivities.

Across the 34 profiles, harmonised natural-spline 20-30 cm bulk density averaged 1.50 g cm^{-3} (SD = 0.05; range 1.42-1.58). The corresponding PCHIP and constant-within-20-40 means were 1.49 and 1.49 g cm^{-3} , respectively.

Candidate gap-filling models were then evaluated against the profile-specific natural-spline values. A linear model using surface bulk density and harmonised deep carbon concentration had the lowest (cross validation - Leave-One-Profile-Out) LOPO RMSE, but its explanatory ability was weak (cross-validated $R^2 = 0.077$), and 23 of the 56 arboretum Tree $\times$ Distance profiles lay outside the reference range for at least one of those predictors. The conservative training mean was therefore selected to avoid unsupported covariate extrapolation.

Table S6c. Leave-one-profile-out prediction of the harmonised 20-30 cm reference values. The training mean was selected for application because predictor-based gains were small and target-domain overlap was incomplete. (OLS: Ordinary Least Square).

Candidate model	Profiles	RMSE (g cm^{-3})	MAE (g cm^{-3})	Bias (g cm^{-3})	Cross-validated R^2
OLS: surface BD + log harmonised deep C	34	0.043	0.035	0.000	0.077
Ridge: surface BD + log harmonised deep C	34	0.044	0.037	-0.002	0.039
OLS: log harmonised deep C	34	0.044	0.037	0.000	0.005
Selected conservative training mean	34	0.046	0.041	0.000	-0.062
OLS: surface BD	34	0.046	0.041	-0.001	-0.072

S2.4 Application to 10-20 and 20-30 cm

Bulk density at 20–30 cm was estimated as the mean of the 34 profile-specific values obtained after harmonising the 0–10, 10–20 and 20–40 cm reference intervals with the selected mass-preserving depth function (1.50 g cm^{-3}). This common value was applied to the arboretum Tree $\times$ Distance profiles and treated as a model-assisted estimate of the 20–30 cm layer mean, rather than as a direct profile-specific measurement. The use of a common layer mean was supported by the weak predictive performance of the profile-specific candidate models and by incomplete overlap between the calibration and application domains. Reference 0–10 cm bulk density ranged from 1.11 to 2.16 g cm^{-3} , whereas arboretum values ranged from 0.80 to 1.30 g cm^{-3} , with 13 of the 56 arboretum Tree $\times$ Distance profiles below the reference minimum. Alternative 20–30 cm estimates obtained using PCHIP interpolation, a constant-within-20–40 cm

assumption and an exponential depth-scaling function were evaluated in the model-form sensitivity analysis reported in Table S8.

Coarse-fragment mass fractions were measured directly at 0–5 and 5–10 cm and converted to volumetric coarse-fragment fractions. Because coarse fragments were not measured below 10 cm and no independent deep-layer calibration was available, the primary calculation carried the measured 0–10 cm volumetric coarse-fragment fraction unchanged through the two deeper layers:

$$CFv_{10-20} = CFv_{20-30} = CFv_{0-10}$$

This was treated as an explicit simplifying assumption rather than as a depth-specific prediction. As a sensitivity analysis, layer-specific coarse-fragment fractions were recalculated from the same measured shallow gravel anchor and the estimated bulk density of each deeper layer. The alternative treatment changed individual 0–30 cm SOC stock estimates by no more than 0.168 Mg C ha⁻¹ (0.255%) and did not alter the statistical conclusions (Table S8).

S2.5 SOC stock calculation and prevention of double correction

Layer-specific SOC stock was calculated as:

$$SOC_stock (Mg\ C\ ha^{-1}) = SOC_40 (g\ kg^{-1}) \times (1 + \theta_h) \times thickness\ (cm) \times \rho_{b,total} (Mg\ m^{-3}) \times (1 - CFv) \times 0.10.$$

The 0.10 conversion factor yields Mg C ha⁻¹ from the stated units. Fixed-depth SOC stock to 30 cm was the sum of the 0-10, 10-20 and 20-30 cm layer stocks. Because $\rho_{b, total}$ included coarse-fragment mass, CFv was accounted for once through (1 - CFv); no additional fine-earth bulk-density transformation was applied (Poeplau et al., 2017).

Equivalent-soil-mass stocks used the smallest cumulative 0-30 cm fine-earth mass among the 56 Tree × Distance profiles as the common reference mass. Carbon and fine-earth mass were accumulated from the surface downward, and only the fraction of the deepest required layer was included to reach the common mass.

S2.6 Calibration and predictive-uncertainty propagation

Uncertainty was propagated in 5,000 draws using two components. First, the 34 unique reference profiles were resampled; the 10-20 cm trimmed ratio and the mean of the profile-specific natural-spline 20-30 cm values were recalculated in every draw. Second, leave-one-

profile-out residuals from the selected 10-20 cm ratio and selected conservative 20-30 cm mean were resampled to represent reference-profile prediction error conditional on the harmonisation. Because the reference data did not identify the correlation of prediction errors among the four distances around the same arboretum tree, two bounding structures were evaluated: one residual per tree and layer shared by all four distances, and independent residuals for every Tree $\times$ Distance profile and layer. A parameter-only bootstrap was retained to distinguish calibration uncertainty from full predictive uncertainty. P-value distributions are reported as sensitivity summaries, not as posterior probabilities of the hypotheses.

Table S7. Primary model-assisted SOC stock results and predictive-uncertainty sensitivity. The point calculation uses the selected layer-specific bulk-density models and constant deep CFv. Ranges are 2.5th-97.5th percentiles across 5,000 draws.

Quantity	Point calculation	Parameter-only bootstrap	Tree-shared prediction error	Profile-independent prediction error
Fixed-depth Distance p	0.0609	0.0608-0.0610	0.0518-0.0720	0.0374-0.1024
Fixed-depth Distance $\times$ Cohort p	0.4254	0.4220-0.4286	0.3820-0.4716	0.3074-0.5599
Fixed-depth 1 m - 8 m estimate (Mg C ha ⁻¹)	14.92	14.84-15.01	14.39-15.46	13.50-16.30
Fixed-depth Holm p for 1 m - 8 m	0.0563	0.0557-0.0571	0.0467-0.0696	0.0349-0.0962
Fixed-depth slope (Mg C ha ⁻¹ per distance doubling)	-4.87	-4.91 to -4.85	-5.05 to -4.71	-5.30 to -4.43
Equivalent-mass Distance p	0.0220	0.0219-0.0222	0.0196-0.0250	0.0169-0.0294
Equivalent-mass Distance $\times$ Cohort p	0.3914	0.3870-0.3954	0.3581-0.4289	0.3255-0.4670
Equivalent-mass 1 m - 8 m estimate (Mg C ha ⁻¹)	16.53	16.43-16.64	15.99-17.06	15.72-17.34
Equivalent-mass Holm p for 1 m - 8 m	0.0189	0.0188-0.0190	0.0166-0.0216	0.0148-0.0245

The fixed-depth categorical result remained near, but above, $\alpha = 0.05$ under the selected point model. Under full predictive-error propagation, 0.7% of tree-shared draws and 19.2% of profile-independent draws gave Distance p < 0.05; the distributions therefore support describing the categorical fixed-depth evidence as uncertain rather than significant. In contrast, the continuous fixed-depth slope and the equivalent-soil-mass Distance effect remained below p = 0.05 in every draw under both error structures. The Distance $\times$ Cohort interaction remained above p = 0.05 in every draw. For the point calculation, the parametric-bootstrap p-values were 0.077 for fixed-depth Distance and 0.413 for Distance $\times$ Cohort. The fixed-depth 1 m versus 8 m contrast had Holm-adjusted p = 0.056 by the tree-level t test and 0.059 by exact sign flip. For equivalent soil mass, the parametric-bootstrap p-values were 0.043 for Distance and 0.392 for Distance $\times$ Cohort; the 1 m versus 8 m contrast had Holm-adjusted p = 0.019 by the t test and 0.027 by exact sign flip.

S2.7 Model-form, coarse-fragment and directly measured surface-stock sensitivities

Table S8. Model-form and coarse-fragment sensitivity of the fixed-depth Distance test. The reference-layer-means benchmark is shown to document the effect of discarding the useful surface-BD information at 10-20 cm; it was not selected because its 10-20 cm cross-validated RMSE was higher than that of the trimmed-ratio model.

Deep-layer model	Deep CFv assumption	Distance F(3,30)	Distance p	Distance × Cohort p
Former surface ratio at both deep layers	Former scaled CFv	2.33	0.0946	0.4644
Natural mass-preserving 20-30 cm mean	Former scaled CFv	2.74	0.0608	0.4255
Primary: natural mass-preserving 20-30 cm mean	Constant measured 0-10 cm CFv	2.74	0.0609	0.4254
PCHIP-harmonised 20-30 cm mean	Constant measured 0-10 cm CFv	2.74	0.0608	0.4257
Constant-within-20-40 mean	Constant measured 0-10 cm CFv	2.74	0.0608	0.4257
Previous smooth-depth-adjusted mean	Constant measured 0-10 cm CFv	2.74	0.0608	0.4257
Reference-layer mean at both deep layers	Constant measured 0-10 cm CFv	3.68	0.0229	0.3117

Table S9. Reference bulk-density profiles from the Waite Rotation Trial (1997) used to calibrate the deep-layer bulk-density gap-filling. Each row is one of the 34 unique profiles retained after deduplication, and each value is the mean of the eastern and western subplots. Bulk density (BD) is in Mg m^{-3} ; total C and total N are in g kg^{-1} ; depth layers are in cm. The complete 1997 dataset and analysis scripts are available from the corresponding author upon reasonable request.

Profile	BD 0-10 (Mg m^{-3})	BD 10-20 (Mg m^{-3})	BD 20-40 (Mg m^{-3})	Total C 0-10 (g kg^{-1})	Total C 10-20 (g kg^{-1})	Total C 20-40 (g kg^{-1})	Total N 0-10 (g kg^{-1})	Total N 10-20 (g kg^{-1})	Total N 20-40 (g kg^{-1})
FWOg 24	1.26	1.52	1.49	12.6	9.7	4.5	0.82	0.67	0.43
FWOg 25	1.33	1.53	1.54	13.3	11.3	4.4	0.88	0.75	0.4
FWOg 26	1.36	1.54	1.56	12.9	4.6	3.1	0.8	0.4	0.33
FWOg 27	1.4	1.55	1.43	12.7	5.4	3.1	0.88	0.43	0.33
FWO 10	1.34	1.45	1.49	14.0	13.1	6.7	0.93	0.98	0.63
FWO 8	1.36	1.47	1.54	10.3	7.8	5.1	0.72	0.67	0.58
FWO 9	1.32	1.54	1.53	13.8	9.8	5.7	0.9	0.77	0.58
FWOg 5	1.14	1.41	1.45	8.6	8.1	5.7	0.65	0.57	0.52
FWOg 6	1.3	1.41	1.51	11.1	8.4	5.3	0.8	0.68	0.58
FWOg 7	1.33	1.47	1.6	11.8	8.5	4.5	0.8	0.68	0.5
FW 34	1.38	1.49	1.55	10.1	7.8	3.5	0.73	0.58	0.35
FW 35	1.34	1.39	1.56	8.9	7.1	3.4	0.63	0.55	0.38
PPFW 30	1.35	1.44	1.53	14.8	9.6	3.6	1.2	0.7	0.4
PPFW 31	1.37	1.48	1.52	15.2	9.6	4.3	1.17	0.75	0.4
PPFW 32	1.25	1.52	1.51	16.5	11.3	5.1	1.23	0.78	0.47
PPFW 33	1.31	1.4	1.46	14.7	10.4	4.7	1.1	0.73	0.45
PPPPWW 11	1.34	1.68	1.51	18.9	12.6	5.8	1.47	1.0	0.63
PPPPWW 12	1.27	1.4	1.43	22.0	15.6	7.8	1.72	1.15	0.7
PPPPWW 13	1.26	1.46	1.47	20.9	14.7	6.4	1.71	1.07	0.6
PPPPWW 14	1.33	1.44	1.46	19.5	10.1	10.1	1.57	0.83	0.83
PPPPWW 15	1.45	1.49	1.56	18.0	9.9	4.3	1.42	0.85	0.53
PPPPWW 16	1.29	1.49	1.52	20.5	12.8	4.6	1.65	0.95	0.53
PPPP 29	1.33	1.46	1.44	20.4	12.8	4.4	1.75	0.9	0.4
PPW 18	1.38	1.53	1.52	17.2	7.7	5.7	1.35	0.7	0.6
PPW 19	1.31	1.41	1.46	16.4	18.0	5.4	1.23	0.93	0.55
PPW 20	1.44	1.41	1.45	15.8	12.7	5.2	1.2	0.98	0.53
PP 1	1.46	1.34	1.46	13.8	10.4	5.4	1.08	0.8	0.53
PP 2	1.38	1.49	1.44	15.6	9.3	5.0	1.25	0.67	0.5
PeWB 21	1.33	1.58	1.49	14.8	9.2	3.8	1.05	0.73	0.45
PeWB 22	1.29	1.44	1.49	15.7	10.7	3.3	1.12	0.77	0.37
PeWB 23	1.3	1.42	1.48	15.0	9.5	4.4	1.05	0.75	0.45
PeW 3	1.11	1.51	1.53	12.5	10.8	5.9	0.87	0.78	0.55
PeW 4	2.16	2.23	1.16	12.7	9.3	4.1	0.91	0.69	0.44
WW 17	1.42	1.46	1.58	14.0	9.5	4.5	1.12	0.8	0.55

The formulas in the companion gravel workbook were reproduced and checked against its saved layer values. The near-identical results for constant and formerly scaled CFv reflect the very small coarse-fragment fractions in these cores. Holding CFv at the measured 0-10 cm value changed any former 0-30 cm profile stock by at most $0.168 \text{ Mg C ha}^{-1}$ (0.255%). Because

both alternatives use the same measured shallow-gravel anchor rather than independent deep samples, this comparison evaluates the effect of the extrapolation rule, not agreement with an external deep-gravel dataset. The constant-depth assumption was selected for transparency, and the audited companion-workbook extrapolation is reported as the sensitivity. The 0-10 cm stock layer used measured bulk density and measured CFv and is independent of deep-layer gap filling. For this directly supported layer, the Distance omnibus test was $F(3,30) = 2.56$, $p = 0.074$ (GG $p = 0.115$); the 1 m stock exceeded the 8 m stock by 9.71 Mg C ha⁻¹ (Holm-adjusted t-test $p = 0.047$; exact sign-flip $p = 0.048$), and the continuous slope was -3.02 Mg C ha⁻¹ per doubling of distance ($p = 0.016$; exact $p = 0.015$). These results are reported as a surface-layer sensitivity rather than as a replacement for the prespecified 0-30 cm analysis.

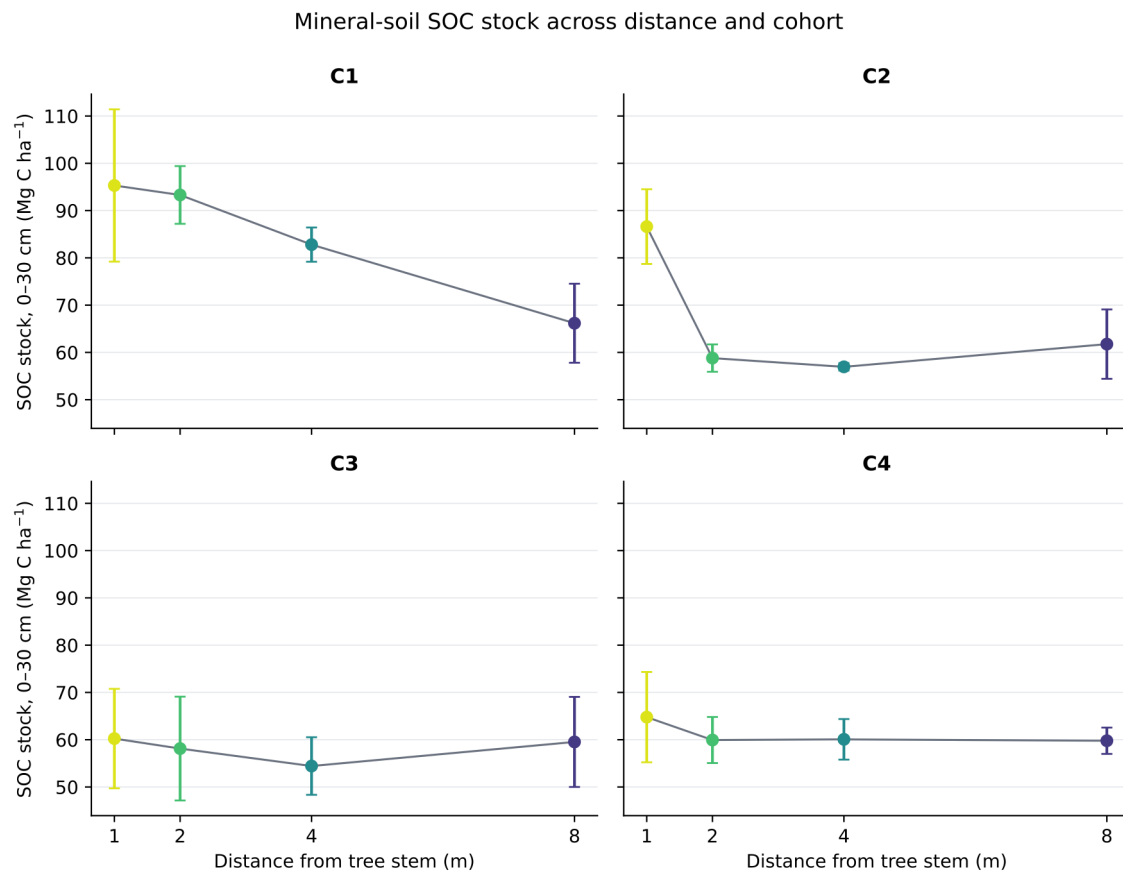


Figure S6. Distance-resolved model-assisted mineral soil SOC stocks integrated to 0–30 cm. Points show cohort means at 1, 2, 4 and 8 m from the tree stem, and error bars show $\pm$ SE among trees within each cohort. Grey lines connect the distance-specific means for visualisation and do not represent fitted regressions. Bulk density below 10 cm was estimated using the validation-selected, depth-harmonised, layer-specific gap-filling procedure, and the measured 0–10 cm coarse-fragment fraction was carried unchanged through both deeper layers in the primary calculation. Statistical inference used tree-level split-plot tests; complete depth-harmonisation, model-form and predictive-uncertainty sensitivities are reported in Supplementary Methods S2.

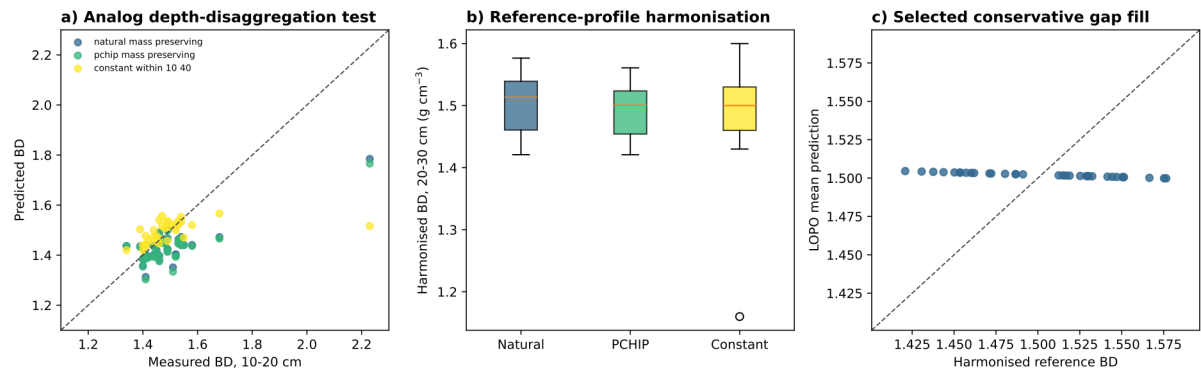


Figure S7. Depth-interval harmonisation and validation of the selected 20–30 cm bulk-density gap filling. (a) Comparison of measured 10–20 cm bulk density with predictions from an analogue depth-disaggregation test in which the 20 cm boundary was concealed within a synthetic 10–40 cm interval. Predictions are shown for the natural mass-preserving, PCHIP mass-preserving and constant-within-interval methods. (b) Distributions of profile-specific 20–30 cm bulk-density estimates obtained using the natural mass-preserving spline, PCHIP and a constant-within-20–40 cm assumption. (c) Leave-one-profile-out predictions from the selected training-mean gap fill plotted against the profile-specific natural-spline estimates. Dashed lines indicate 1:1 agreement. Complete numerical comparisons and influence sensitivities are reported in Tables S6b–S8.

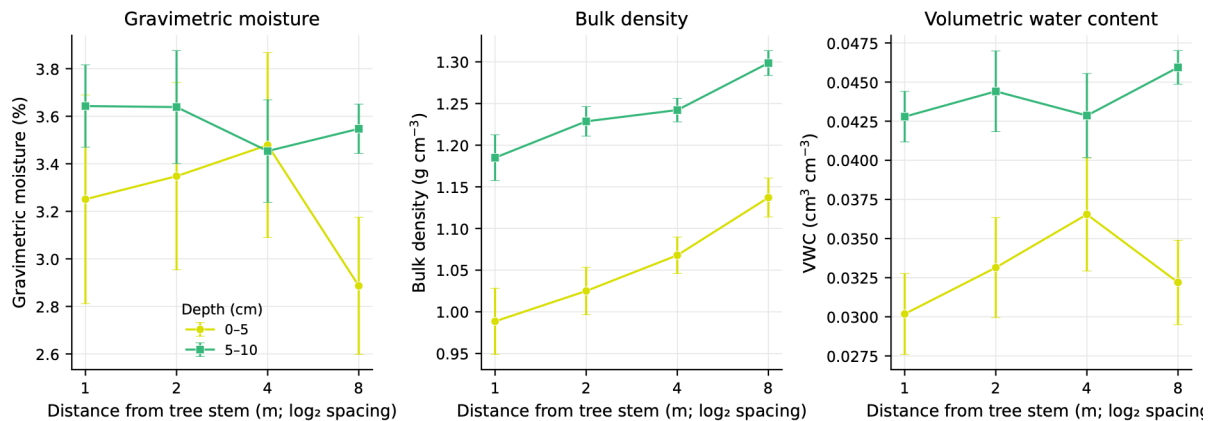


Figure S8. Gravimetric moisture (left), bulk density (centre) and volumetric water content (right) across distance from the tree stem, for the directly measured 0–5 and 5–10 cm layers. Points show tree-level means $\pm$ SE ($n = 14$ trees). Volumetric water content ($VWC = \text{gravimetric moisture} \times \text{bulk density} / \text{water density}$) shows no radial trend: the lower near-tree bulk density is offset by gravimetric moisture that does not decline toward the stem.

S2.8 Reproducibility records and interpretation boundary

The deduplicated reference profiles, cumulative-integral harmonisation outputs, analogue depth-disaggregation validation, influence sensitivity, candidate-model leave-one-profile-out predictions, arboretum applicability-domain audit, selected-model point predictions, calibration and residual-bootstrap draws, fixed-depth and equivalent-soil-mass sensitivity tables, coarse-fragment audit, Python scripts and environment metadata are retained by the authors and are available from the corresponding author upon reasonable request. The depth-harmonised procedure resolves the mismatch between the available 20–40 cm reference interval and the required 20–30 cm study layer without treating them as equivalent, but it does not

replace direct deep intact-core measurements. The 20-30 cm bulk density remains a reference-calibrated expected value, and the resulting 0-30 cm fixed-depth, equivalent-mass and footprint-weighted SOC stocks should be interpreted as model-assisted estimates. Direct deep-layer bulk density and coarse-fragment measurements would be required to convert these estimates into fully measurement-based stocks (Goidts et al., 2009; Walter et al., 2016; Xu et al., 2015).

References

All literature references cited in this supplement are compiled within the reference list of the main manuscript file.