Supplementary Material

Multi-paradigm benchmarking for integrated soil quality assessment under conservation management in a tropical Ultisol

This document contains four supplementary tables referenced in the main manuscript.

Table S1: PCA eigenvalues, variance explained, and component weights for the ANOVA/MDS paradigm (34 indicators). Only the eight retained components (eigenvalue > 1) are shown; together they explain 83.56% of total variance.

Component	Eigenvalue	Var. expl. (%)	Cumul. (%)	Weight[a]
PC1	8.230	24.21	24.21	0.310
PC2	5.695	16.75	40.96	0.215
PC3	3.284	9.66	50.61	0.124
PC4	2.934	8.63	59.24	0.111
PC5	2.566	7.55	66.79	0.097
PC6	2.435	7.16	73.95	0.092
PC7	1.794	5.28	79.23	0.068
PC8	1.471	4.33	83.56	0.056

[a]Weight = eigenvalue_i / $\sum_{i=1}^{8}\text{eigenvalue}_i$; used in the additive $\text{SQI}_{\text{ANOVA}}$ computation.

Table S2: Complete expert rule base of the hierarchical Mamdani fuzzy inference system (IPACSA). Membership functions (MFs): L = Low, M = Medium, H = High; 0 = don't care. Connection: AND (minimum t-norm). Defuzzification: centroid.

Sub-index (inputs)	Rule	Antecedent pattern	Output	Weight
COS – Soil Conservation (7 inputs: DMG, DMP, RMP, Ds, Na, ICV, COS_{stock})				
	E1	M, M, L, L, L, H, H	H	1.0
	E2	M, M, M, M, M, M, M	M	1.0
	E3	L, L, H, H, H, L, L	L	1.0
	F1–F21	Main-effect fallback[a]	L/M/H	0.1
ARQ – Plant Architecture (6 inputs: ALT, DESP, CESP, NESP, NESPC, PESP)				
	E1	H, H, H, H, H, H	H	1.0
	E2	M, M, M, M, M, M	M	1.0
	E3	L, L, L, L, L, L	L	1.0
	F1–F18	Main-effect fallback[a]	L/M/H	0.1
STAND – Plant Stand (1 input: STAND)				
	E1	H	H	1.0
	E2	M	M	1.0
	E3	L	L	1.0
IPACSA – Final aggregation (4 inputs: COS, ARQ, STAND, PROD)				
	E1	M, H, M, H	H	1.20
	E2	M, M, M, M	M	1.00
	E3	L, L, L, L	L	0.80
	E4	H, M, M, M	H	1.10
	E5	M, H, M, H	H	1.10
	E6	M, M, H, M	H	1.10
	E7	M, H, M, M	H	1.15
	F1–F12	Main-effect fallback[a]	L/M/H	0.1

[a]Main-effect fallback rules: for each input variable k, three rules are generated where only the k-th input is set to L, M, or H (all others = 0, i.e., don't care), mapping directly to the corresponding output level. These rules (weight = 0.1) prevent NaN outputs in regions of the input space not covered by expert rules.

Table S3: PLS-SEM structural model: path coefficients, bootstrap confidence intervals, and model fit (five constructs, 25 manifest variables after VIF screening).

Path	γ	Boot. Mean	Boot. SD	t	95% CI	p
PHYSICAL $\rightarrow$ YIELD	-0.122	-0.013	0.160	0.76	$[-0.28, 0.25]$	0.992
CHEMICAL $\rightarrow$ YIELD	0.097	0.104	0.102	0.95	$[-0.14, 0.28]$	0.264
MICRO $\rightarrow$ YIELD	0.175	0.119	0.127	1.39	$[-0.22, 0.28]$	0.284
AGRO $\rightarrow$ YIELD	0.622	0.608	0.051	12.30	[0.50, 0.70]	<0.001
$R^2 = 0.690$; Adj. $R^2 = 0.678$; Bootstrap: 500 resamples, seed = 42.						

Table S4: PLS-SEM outer weights (Mode B, formative) and bootstrap significance for 25 retained manifest variables.

Construct	Indicator	Weight	Boot. SD	t	95% CI	p
PHYSICAL	DMG	0.068	0.192	0.35	[−0.29, 0.41]	0.792
	COS	0.086	0.217	0.40	[−0.39, 0.47]	0.784
	Na	0.153	0.216	0.71	[−0.36, 0.43]	0.864
	ICV	0.031	0.196	0.16	[−0.34, 0.44]	0.888
	MI	−0.301	0.315	0.95	[−0.50, 0.60]	0.800
	MA	0.684	0.691	0.99	[−1.09, 1.17]	0.924
	VIB	0.178	0.268	0.66	[−0.48, 0.55]	0.900
	RP	0.314	0.374	0.84	[−0.66, 0.68]	0.980
	AD/PT	1.200	1.108	1.08	[−1.41, 1.47]	0.924
CHEMICAL	pH	−0.365	0.206	1.77	[−0.61, 0.42]	0.092
	P	−0.508	0.263	1.93	[−0.75, 0.52]	0.088
	K	−0.044	0.242	0.18	[−0.51, 0.44]	0.936
	Mg	0.096	0.196	0.49	[−0.32, 0.44]	0.640
	CEC	−0.197	0.210	0.94	[−0.51, 0.46]	0.316
	EstN	0.618	0.330	1.87	[−0.51, 1.06]	0.088
MICRO	qMIC	−0.580	0.445	1.30	[−0.98, 0.80]	0.276
	qCO_2	0.357	0.295	1.21	[−0.43, 0.78]	0.276
	CCO_2	−0.246	0.363	0.68	[−0.72, 0.75]	0.656
	NMIC	0.927	0.596	1.56	[−0.90, 1.05]	0.276
AGRO	ALT	0.601	0.089	6.74	[0.40, 0.76]	<0.001
	DESP	0.124	0.087	1.43	[−0.06, 0.30]	0.172
	CESP	0.337	0.081	4.16	[0.17, 0.48]	<0.001
	STAND	0.529	0.108	4.90	[0.32, 0.72]	0.004
YIELD	NESP	0.534	0.051	10.46	[0.44, 0.62]	<0.001
	PESP	0.643	0.052	12.34	[0.56, 0.76]	<0.001

VIF screening (threshold = 5) excluded Ca (VIF = 26.5), MOS (VIF = 9.8), MBC (VIF = 23.9), and GY (VIF = 20.1). YIELD construct estimated with Mode A (reflective); all others with Mode B (formative). Bootstrap: 500 resamples.

Table S5: Random Forest variable importance (34-indicator model, 500 trees, $m_{\text{try}} = 11$). Variables ranked by permutation importance (%IncMSE).

Rank	Indicator	Domain	%IncMSE	IncNodePurity
1	NESP	Yield	20.22	30.46
2	STAND	Agronomic	19.49	23.51
3	NESPC	Yield	13.19	8.44
4	PROD	Yield	13.15	8.02
5	PESP	Yield	12.51	8.26
6	DESP	Agronomic	12.33	9.38
7	ICV	Physical	11.19	5.95
8	CESP	Agronomic	11.13	5.27
9	P	Chemical	8.04	3.01
10	Ca	Chemical	7.63	1.32
11	SOM	Chemical	7.35	1.68
12	CEC	Chemical	6.68	1.48
13	K	Chemical	6.64	1.81
14	MI	Physical	6.47	1.16
15	pH	Chemical	6.38	1.35
16	EstN	Chemical	6.11	1.36
17	Al	Chemical	6.09	1.23
18	qMIC	Microbiol.	5.89	1.11
19	MBC	Microbiol.	5.64	0.97
20	AD/PT	Physical	5.51	1.20
21	VIB	Physical	5.39	1.16
22	BR	Microbiol.	5.23	1.12
23	MBN	Microbiol.	5.09	1.15
24	RP	Physical	4.96	0.86
25	Mg	Chemical	4.68	0.60
26	RMP	Physical	4.53	1.29
27	qCO_2	Microbiol.	4.38	0.80
28	MA	Physical	4.31	1.00
29	COS	Physical	4.18	1.78
30	GMD	Physical	3.87	1.10
31	ALT	Agronomic	3.40	1.39
32	WMD	Physical	3.30	1.56
33	Na	Physical	2.97	1.10
34	BD	Physical	2.95	1.05

%IncMSE, percentage increase in mean squared error upon permutation of the variable; IncNodePurity, total decrease in node impurity (residual sum of squares). Domain abbreviations: Microbiol. = Microbiological.