



Hungarian University of Agriculture and Life Sciences.

From Legacy Data to a National Knowledge System:
Development of a Namibian Soil Information System, National Soil Map,
and Spatial Predictions of Soil Properties

Marina Elda Coetzee
Gödöllő, Hungary
2026

The PhD School

Name of Doctoral School: Natural Sciences

Discipline: Soil Science

Head: Prof. Erika Csakiné Michéli, MHAS
Professor, DSc
MATE, Institute of Environmental Sciences
Department of Soil Science

Supervisors: Prof. Erika Csakiné Michéli, MHAS
Professor, DSc
MATE, Institute of Environmental Sciences
Department of Soil Science

Dr. Ádám Csorba
Associate Professor, PhD
MATE, Institute of Environmental Sciences
Department of Soil Science

.....
Approval of the Head of Doctoral School

.....
Approval of the Supervisors

DECLARATION

This dissertation is my original work and has not been presented for a degree in any other university. No part of this dissertation may be reproduced without prior permission of the author and/or the Hungarian University of Agriculture and Life Sciences.

..... Date:

DEDICATION

This work is dedicated to my late father, Bertus, mother, Bertha, and sister, Elmarie, whose love, encouragement, and belief in me continue to inspire me. Although they could not share in the completion of this work, their inspiration and influence have remained with me throughout, and this achievement is also part of their legacy.

TABLE OF CONTENTS

DECLARATION	i
DEDICATION	ii
TABLE OF CONTENTS	iii
TABLES	vii
FIGURES	viii
ACRONYMS AND ABBREVIATIONS	x
1. INTRODUCTION	1
1.1 Context	1
1.2 Problem statement	2
2. AIM AND OBJECTIVES	3
2.1 Research aim	3
2.2 Specific objectives	3
2.3 Study approach and hypotheses	3
2.4 Dissertation structure	4
3. CONTEXT AND LITERATURE REVIEW	6
3.1 Environmental conditions and land use in the Republic of Namibia	6
3.2. Major dryland pedogenesis processes and the influence of lithology	10
3.3 Main soil types of Namibia	13
3.4 Global and African soil information and mapping initiatives	16
3.5 Soil information systems	17
3.6 Legacy soil data	19
3.6.1 Data archaeology	20
3.6.2 Data rescue	20
3.6.3 Data renewal	21
3.7 Conventional soil mapping (CSM)	21
3.7.1 Conceptual CSM framework	21
3.7.2 Methodological considerations in synthesising a national map from legacy surveys	22
3.7.3 Strengths and limitations of CSM	23
3.8 Digital soil mapping (DSM)	24
3.8.1 Conceptual DSM frameworks	24
3.8.2 Machine-learning approaches in predictive soil mapping	25
3.8.3 Methodological considerations in DSM	26
3.8.3.1 Quality of soil data	26
3.7.3.2 Choice of machine learning algorithm	26
3.8.3.3 Choice of modelling environment	27
3.8.3.4 Covariate selection	28
3.8.3.5 Model tuning, training and validation	28
3.8.3.6 Performance metrics	29
3.8.3.7 Uncertainty estimation	30
3.8.4 Strengths, limitations, knowledge gaps, and current debates	30
4. METHODS	33
4.1 Overview	33
4.2 Conceptualisation and design of the Namibian Soil Information System (NamSIS)	34
4.2.1 Assessment of user needs	34
4.2.2 Existing systems and standards	34

4.2.3	System architecture concept	35
4.2.4	Alignment with fair principles	35
4.3	Legacy data collection, rescue and revitalisation	35
4.3.1	Collection and assessment of legacy materials (data archaeology)	35
4.3.2	Digitisation and documenting (data rescue)	36
4.3.3	Development of the Namibian Soil Profile Database (data renewal)	36
4.3.3.1	Source data	37
4.3.3.2	Data entry, cleaning and quality control	40
4.3.3.3	Data enhancement and integration	41
4.4	Production of a conventional national soil map of Namibia	42
4.4.1	The basemap	42
4.4.2	Supplementary legacy soil maps and databases	44
4.4.3	Ancillary information	45
4.4.4	Completion of previously unmapped regions	49
4.4.5	Integration — the National Soil Map of Namibia	49
4.4.6	Map unit design and classification	50
4.4.7	Quality control and validation	51
4.5	Predictive spatial modelling	52
4.5.1	Overview of DSM workflow	52
4.5.2	Choice of modelling platform and algorithm	53
4.5.3	Soil data	53
4.5.4	Environmental covariates	54
4.5.5	Covariate selection	55
4.5.6	Hyperparameter tuning, model training and validation	55
4.5.7	Spatial predictions and uncertainty estimation	56
4.5.8	Pedological plausibility evaluation	56
4.5.9	External validation	57
4.5.10	Mapping pipeline	57
4.6	Integration — the Namibian Soil Information System (NamSIS)	57
5.	RESULTS AND DISCUSSION	59
5.1	Legacy data preservation and documentation	59
5.1.1	Original point (profile) data archive	59
5.1.2	Analytical methods	59
5.1.3	Documents archive	60
5.1.4	Photographs archive	60
5.1.5	Original maps archive	61
5.2	The Namibian Soil Profile Database (NSPD)	61
5.2.1	Database structure	61
5.2.2	Data volume and dataset composition	62
5.2.3	Pedogenic horizons, fixed depth layers, and depth-harmonised data	63
5.3	The National Soil Map of Namibia (NSMN)	63
5.3.1	Overview	63
5.3.2	Distribution of RSGs, environmental controls, and implications for land use and management	66
5.3.3	Derived thematic maps	70
5.3.3.1	Categorical maps of individual soil classes	70
5.3.3.2	Categorical maps of selected soil properties	71
5.3.3.3	Terrain attributes	73
5.3.4	Advances and strengths of the national soil map	76
5.3.5	Limitations and sources of uncertainty	77
5.4	Digital soil property maps of Namibia	78
5.4.1	Soil data	78
5.4.2	Feature selection	80

5.4.3	Model optimisation	81
5.4.4	Model performance (quantitative results)	81
5.4.5	Variable importance and environmental controls on pedogenesis	83
5.4.6	Spatial predictions and pedological plausibility assessment	85
5.4.7	Prediction uncertainty	94
5.4.8	External validation	97
5.4.9	Limitations and uncertainty	98
5.4.10	Complementary roles of conventional and digital soil mapping	99
5.5	The Namibian Soil Information System (NamSIS)	99
5.5.1	System structure	99
5.5.2	Geographical information system	99
5.5.3	Digital soil modelling scripts	100
5.5.4	Image products	101
5.5.5	Metadata	101
5.5.6	System design strengths and limitations	101
5.5.7	Practical considerations and data access	102
6.	CONCLUSIONS AND RECOMMENDATIONS	104
6.1	Achieving the study's objectives	104
6.1.1	Development of a national soil information framework	104
6.1.2	Legacy data recovery and consolidation	104
6.1.3	Conventional soil mapping	104
6.1.4	Digital soil modelling	104
6.2	Conceptual, methodological, and scientific contributions	105
6.3	Hypothesis support	105
6.3.1	Hypothesis H1 — Data integration	105
6.3.2	Hypothesis H2 — Pedogenic-environmental relationships	106
6.3.3	Hypothesis H3 — Conventional-digital soil mapping convergence	106
6.3.4	Overarching hypothesis	107
6.4	Recommendations for future work	107
6.4.1	Improving the observational basis	107
6.4.2	Enhancing the national soil map	107
6.4.3	Advancing digital soil mapping	108
6.4.4	Further development of NamSIS and user access	108
6.4.5	Institutional issues	109
6.5	Final remark	110
7.	NEW SCIENTIFIC RESULTS	111
	SUMMARY	112
	APPENDICES	116
A.	References	116
B.	Environmental covariates assembled in the predictor stack	134
C.	Thumbnail maps of all environmental predictors	139
D.	Digital soil mapping pipeline	143
E.	Distribution of the NSPD's soil data points per dataset	147
F.	NSPD MS Access tables and fields	150
G.	NSPD MS Excel tables and fields	154
H.	Soil typological units identified in Namibia	158
I.	Spatial distribution of soil observations used for modelling, per soil property	162
J.	Temporal distribution of soil data	163
K.	Boxplots of depth-harmonised soil property data	164

L.	Histograms of depth-harmonised legacy soil property data	166
M.	Environmental variables included in RF models — all properties and depths	168
N.	Random Forests hyperparameters optimised for each soil property and depth interval	169
O.	Descriptive statistics of performance metrics across all bootstrap iterations	170
P.	Spatial maps of prediction uncertainty for all soil properties and depth intervals	172
Q.	PI90 maps for all soil properties and depth intervals	188
R.	Prediction interval coverage probability (PICP) plots for all soil properties and depth intervals	191
S.	Scatterplots with validation statistics from external validation comparing predictions from this study, iSDA, and SoilGrids250m with independent Soils4Africa observations	194
T.	Summary of model performance metrics reported in selected digital soil mapping studies	197
U.	NamSIS datasets and scripts	203
ACKNOWLEDGEMENTS		205

TABLES

Table 1	Soil properties selected for modelling	53
Table 2	Number and types of photos by dataset in the Namibian Soil Profile Database (NSPD)	60
Table 3	Current datasets in the Namibian Soil Profile Database (NSPD) and their respective number of sites (profiles, augerings, or sampling points)	62
Table 4	The current number of laboratory records per soil property in the Namibian Soil Profile Database (NSPD)	63
Table 5	Descriptive statistics of depth-harmonised soil properties across Namibia for three standard depth intervals (0–30, 30–60, and 60–100 cm)	79
Table 6	Frequency (%) with which environmental covariates were included in all Random Forests models.	81
Table 7	Mean model performance metrics (RMSE and R^2) across 20 bootstrap iterations for each soil property and depth interval	82
Table 8	Prediction Interval Coverage Probability (PICP) for each soil property and depth interval, to assess the robustness of the 90% prediction intervals	94
Table 9	External validation metrics (RMSE and R^2) comparing predictions from this study with iSDA and SoilGrids250m V2 products, against independent Soils4Africa observations.	97

FIGURES

Figure 1	Elevation of Namibia above mean sea level	6
Figure 2	Main physiographic regions of Namibia.	6
Figure 3	Landscapes of Namibia (Atlas of Namibia Team, 2022)	7
Figure 4	Broad lithological classes of Namibia (derived from Geological Survey of Namibia, n.d.) (mapping incomplete in northeast)	7
Figure 5	Mean annual rainfall of Namibia (Atlas of Namibia Team, 2022)	8
Figure 6	Aridity index of Namibia (Atlas of Namibia Team, 2022)	8
Figure 7	Sub-biomes (based on Strohbach et al., 2025)	9
Figure 8	Major farming systems of Namibia (Atlas of Namibia Team, 2022).	10
Figure 9	The SCORPAN model (McBratney et al, 2003; Yigini et al, 2018).	25
Figure 10	Workflow of the study — Development of a Namibian soil information system, national soil map, and spatial predictions of soil properties	33
Figure 11	Datasets contributing to the Namibian Soil Profile Database	38
Figure 12	Locations of the four NSS surveys (ICC and MAWRD, 2000).	42
Figure 13	Preliminary Soil Map of Namibia (Coetzee, 2001).	43
Figure 14	Land Systems of the 2021 Agro-Ecological Zones Map of Namibia (Van der Waal et al., 2021a, 2021c)	48
Figure 15	Differences in the colour palettes of the Soil Atlas of Africa (Jones et al., 2013) and the new National Soil Map of Namibia	51
Figure 16	Digital soil mapping workflow.	52
Figure 17	Percentage of samples per laboratory in the Namibian Soil Profile Database (NSPD).	60
Figure 18	Relational structure of the Namibian Soil Profile Database.	61
Figure 19	Distribution of soil point data in (a) the Namibian Soil Profile Database (NSPD) and (b) the 1998-2000 National Soil Survey of Namibia.	62
Figure 20	The new National Soil Map of Namibia — dominant World Reference Base (WRB) Reference Soil Groups (RSGs)	64
Figure 21	The new National Soil Map of Namibia (NSMN) — outlines of all soil map units	65
Figure 22	Excerpts from the new National Soil Map of Namibia showing the coverages of previously unmapped areas unmapped in the NSS 1M map (ICC & MAWRD, 2000): (a) Skeleton Coast National Park, (b) Tsau I Khaeb National Park, and (c) Etosha National Park	66
Figure 23	Excerpts from the new National Soil map of Namibia ((a) and (c)) and the NSS 1M Map ((b) and (d)) to illustrate the increase in spatial detail: (a) and (b) northeastern Namibia; (c) and (d) north-central Namibia	66
Figure 24	Distribution of major WRB RSGs to indicate where each one is spatially dominant ($\geq 50\%$) (dark blue), associated (25–50%) (medium blue) or included (5–25%) (light blue). (Full resolution maps can be viewed and downloaded at https://soilsofnamibia.com/maps.html)	70
Figure 25	Distribution of selected soil properties and WRB diagnostic features, to indicate where each is spatially dominant ($\geq 50\%$) (dark green), associated (25–50%) (medium green) or included (5–25%) (light green). (Full resolution maps can be viewed and downloaded at https://soilsofnamibia.com/properties.html)	71
Figure 26	Slope Map of Namibia (GAEZ classes)	73
Figure 27	Percentage of each SMU with slopes steeper than 8% in Namibia	73
Figure 28	Aspect maps of Namibia: (a) Dune orientation highlighted in the Sossusvlei area and part of the Namib Sand Sea; (b) Subtle omiramba drainage patterns in the Epukiro-Eiseb area	74
Figure 29	Combined aspect-slope maps of Namibia: (a) Inselbergs in the Namib, river gorges, the Schwarzrand scarp, and the abrupt transition from the rugged escarpment to the coastal plain and Namib Sand Sea; (b) Brukkaros Mountain, the Weissrand scarp, river valleys cutting through the Kalahari Sands, and the transition from the relatively level interior plateau to the incised escarpment	74
Figure 30	Terrain Ruggedness Index (TRI): (a) Namibia; (b) Marienfluss–Epupa area; (c) Brandberg–Messum–Ugab River area; (d) Windhoek Valley, Khomas Hochland, and Auas Mountains; (e) Sossusvlei area, highlighting star and transverse dunes, the valley of the Tsauchab River and terminal playas; (f) Aussenkehr area and the confluence of the Fish and Orange Rivers	75
Figure 31	Ruggedness: Percentage of (a) flat; (b) level; (c) nearly level; (d) slightly rugged; (e) moderately rugged; and (f) highly rugged land per soil map unit in Namibia.	75, 76
Figure 32	Topographic Position Index (TPI): (a) Namibia; (b) Brandberg–Ugab River area; and (c) Schwarzrand.	76
Figure 33	Relative importance of environmental covariates in Random Forests models, aggregated across soil	84

	properties and depth intervals, based on RF variable importance scores. The ten most influential covariates are ranked in decreasing order from 1 (most important) to 10. Other variables, included in modes but of lower importance, are shown in light yellow ('included')	
Figure 34	National-scale spatial predictions of bulk density (BD) (top) and base saturation (BS) (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at https://soilsofnamibia.com/properties.html)	86
Figure 35	National-scale spatial predictions of pH (H ₂ O) (top) and cation exchange capacity (CEC) (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at https://soilsofnamibia.com/properties.html)	87
Figure 36	National-scale spatial predictions of extractable calcium (Ca) (top) and magnesium (Mg) (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at https://soilsofnamibia.com/properties.html)	88
Figure 37	National-scale spatial predictions of extractable potassium (K) (top) and sodium (Na) (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at https://soilsofnamibia.com/properties.html)	89
Figure 38	National-scale spatial predictions of electrical conductivity (EC) measured in a 1:2.5 soil:water extract (top) and in a saturated paste extract (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at https://soilsofnamibia.com/properties.html)	90
Figure 39	National-scale spatial predictions organic carbon (OC) (top) and plant-available phosphorus (P) (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at https://soilsofnamibia.com/properties.html)	91
Figure 40	National-scale spatial predictions of total nitrogen (N) (top) and sand content (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at https://soilsofnamibia.com/properties.html)	92
Figure 41	National-scale spatial predictions of silt (top) and clay content (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at https://soilsofnamibia.com/properties.html)	93
Figure 42	Spatial distribution of prediction uncertainty for pH (H ₂ O) in Namibia. For each depth interval (0–30, 30–60, and 60–100 cm) (left to right), maps show the 5th (top row) and 95th percentile (bottom row) predictions. (Full resolution maps can be viewed and downloaded at https://soilsofnamibia.com/properties.html)	95
Figure 43	Spatial distribution of prediction uncertainty for pH (H ₂ O) in Namibia. For each depth interval (0–30, 30–60, and 60–100 cm) (left to right), maps show the mean prediction (top row) and 90% prediction interval (PI90) (bottom row). (Full resolution maps can be viewed and downloaded at https://soilsofnamibia.com/properties.html)	96
Figure 44	Components of the Namibian Soil Information System (NamSIS)	100

ACRONYMS AND ABBREVIATIONS

%	percent
µm	micrometer
µS cm ⁻¹	MicroSiemens per Centimeter
AEZ	Agro-Ecological Zones
AfSIS	Africa Soil Information Service
AfSoilGrids250m	African Soil Property Prediction Maps at 250 m resolution
AgricLab	Agriculture Laboratory (of the Namibian Ministry of Agriculture)
AgriLASA	Agri Laboratory Association of Southern Africa
ALOS	Advanced Land Observing Satellite
AMP	(laboratory of the) University of Pretoria's Centre for Microbial Ecology and Genomics
API	Application Programming Interface
ARC-ISCW	Agricultural Research Council (of South Africa) – Institute for Soil, Climate and Water
ArcGIS	(trademarked, proprietary geographical information system software, from ESRI)
ASPP	Average Seasonal Plant Productivity
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer (instrument aboard Terra satellite)
AUSO	African Union Soil Observatory
BD	Bulk Density
BIOTA	Biodiversity Monitoring Transect Analysis in Africa (Project)
BS	Base Saturation
Ca	Calcium
CaCl ₂	Calcium Chloride
CaCO ₃	Calcium Carbonate
CC	Carrying Capacity
CCC	Lin's Concordance Correlation Coefficient
CD	Compact Disc
CEC	Cation Exchange Capacity
CEC _{effective}	Effective Cation Exchange Capacity
CHILI	Continuous Heat-Insolation Load Index
CLORPT	Climate, Landform, Organisms, Parent Material, Time (Model)
cmol(+) kg ⁻¹	Centimol of Charge per Kilogram
CSM	Conventional Soil Mapping
CSV	Comma-Separated Values
DEM	Digital Elevation Model
DFG	Deutsche Forschungsgemeinschaft (German Research Foundation)
DOI	Digital Object Identifier
DSM	Digital Soil Mapping / Digital Soil Model(s)
DSMW	Digital Soil Map of the World
E	East (in coordinates)
EC	European Commission
EC	Electrical Conductivity
EIS	Environmental Information Service of Namibia
ENSO	El Niño Southern Oscillation
EPSG	European Petroleum Survey Group (organisation that maintains a geodetic parameter database with standard EPSG codes for coordinate systems, datums spheroids, units, etc.)
ESA	European Space Agency
ESDAC	European Soil Data Centre (hosted by JRC)
ESRI	Earth Systems Research Institute
EU	European Union
EVI	Enhanced Vegetation Index
FAIR	Findable, Accessible, Reusable and Interoperable
FAO	Food and Agriculture Organization (of the United Nations)
GAI	Global Aridity Index
GDAL	Geospatial Data Abstraction Library
GEE	Google Earth Engine (cloud computing platform)
GeoTIFF	TIFF with embedded georeferencing information (image format)
GIS	Geographical Information System
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit (German International Cooperation)

GLOSIS	Global Soil Information System
GLOSOLAN	Global Soil laboratory Network of the FAO
GSN	Geological Survey of Namibia
GSP	Global Soil Partnership
GTZ	Deutsche Gesellschaft für Technische Zusammenarbeit (German Technical Cooperation)
H ₂ O	Water
HWSO	Harmonised World Soil Database
HydroSHEDS	Hydrological data and maps based on SHuttle Elevation Derivatives at multiple Scales
ICC	Institut Cartogràfic de Catalunya (Cartographic Institute of Catalonia)
IIASA	International Institute for Applied Systems Analysis
INSI	International Network of Soil Information Institutions
iSDA	Innovative Solutions for Decision Agriculture
ISO	International Standards Organisation
ISRIC	International Soil Reference and Information Center
ISSCAS	Institute of Soil Science, Chinese Academy of Sciences
ITCZ	Intertropical Convergence Zone
IUCN	International Union for Conservation of Nature
IUSS	International Union of Soil Scientists
IWRMP	Integrated Water Resources Management Plan
JRC	Joint Research Centre (of the European Commission)
JPEG	Joint Photographic Experts Group (image file format)
JPL	Jet Propulsion Laboratory
JSON	JavaScript Object Notation (data exchange format)
K	Potassium
kg dm ⁻³	Kilogram per Cubic Decimeter
kNDVI	Kernel Normalised Difference Vegetation Index
LandPKS	Land Potential Knowledge System (Project)
LDN	Land Degradation Neutrality (Project)
LERIS	Landscape Ecological Risk Information System
MAWF	Ministry of Agriculture, Water and Forestry (2000-2020)
MAWF-AEZ	Agro-Ecological Zoning Programme of the Directorate of Agricultural Research and Training, Ministry of Agriculture, Water and Forestry
MAWFLR	Ministry of Agriculture, Water, Fisheries and Land Reform (since 2025)
MAWLR	Ministry of Agriculture, Water and Land Reform (2020-2025)
MAWRD	Ministry of Agriculture, Water and Rural Development (1991-2000)
MEFT	Ministry of Environment, Forestry and Tourism
meq	Milli-Equivalent
MET	Ministry of Environment and Tourism
Mg	Magnesium
mg kg ⁻¹	Milligram per Kilogram
MIR	Mid-Infrared (spectroscopy)
ML	Machine Learning
MoRAP	Missouri Resources Assessment Partnership
MS Access	Microsoft Access (database management system / software)
MS Excel	Microsoft Excel (spreadsheet editing software)
MSAVI	Modifiedt Soil Adjusted Vegetation Index
N	Nitrogen
Na	Sodium
NA	(laboratory of the) Namibian Ministry of Agriculture (= AgricLab)
NamSIS	Namibian Soil Information System
NamSOTER	Namibia Soil and Terrain Database
NARIS	Namibian Agricultural Resources Information System
NASA	National Aeronautics and Space Administration
NAW	Neighbourhood Analysis Window
NDVI	Normalised Difference Vegetation Index
NIR	Near-Infrared (spectroscopy)
NRCS	Natural Resources Conservation Service (of the USDA)
NSA	National Statistics Agency
NSDI	Namibian Spatial Data Infrastructure
NSMN	National Soil Map of Namibia
NSPD	Namibian Soil Profile Database
NSS	National Soil Survey Phase I, carried out by MAWRD and ICC in 1998-2000
NSS 1M	Preliminary Soil Map of Namibia (1:1 million scale) produced in 2000

NSS OKA	National Soil Survey Phase I Soil Map of the Okavango River Terrace (1:100,000 scale) produced in 2000
NSS OWA	National Soil Survey Phase I Soil Map of the Central Cuvelai Basin (1:100,000 scale) produced in 2000
NSS TWO	National Soil Survey Phase I Soil Map of Northeast Namibia (1:250,000 scale) produced in 2000
NUST	Namibia University of Science and Technology
OC	Organic Carbon
ORI	(laboratory of the) University of Botswana's Okavango Research Institute
P	Phosphorus
PDF	Portable Document Format
PET	Potential Evapotranspiration
pH (H ₂ O)	pH (1:2.5 Soil:Water)
PhD	Doctor of Philosophy
PI90	90% Prediction Interval
PICP	Prediction Interval Coverage Probability
PNG	Portable Network Graphics (raster image file format)
PRID	Profile Identification
QGIS	Quantum Geographic Information System (free, open-source, cross-platform desktop geographical information system software)
R	R (programming language for statistical computing and data visualisation)
R ²	Coefficient of Determination
RF	Random Forests (machine learning algorithm)
RMSE	Root Mean Squared Error
RSG	Reference Soil Group
S	South (in coordinates); Soil in models
S4A	Soils4Africa (Project)
SASSCAL	Southern African Science Service Centre for Climate Change and Adaptive Land Management (Programme)
SAVI	Soil Adjusted Vegetation Index
SCIONA	Co-designing conservation technologies for Iona-Skeleton Coast Transfrontier Conservation Area (Project)
SCORPAN	Soil attributes, Climatic properties, biOtic activity, terrain features/Relief, Parent material, soil Age, Neighbourhood context (Model)
SIS	Soil Information System
SMU	Soil Map Unit
SOC	Soil Organic Carbon
SOM	Soil Organic Matter
SOP	Standard Operating Procedure
SoilGrids250	Global soil property prediction and soil classes maps at 250m resolution
SoilGrids1km	Global soil property prediction and soil classes maps at 1km resolution
SOTER	Soil and Terrain Database
SOTERSAF	Soil and Terrain Database for Southern Africa
SOTWIS	SOTER and WISE-based databases of harmonised soil parameter estimates
SRTM	Shuttle Radar Topography Mission
STEP-AWBH	Soil, Topographic, Ecological, Parent Material, Atmospheric, Water, Biotic and Human-induced Properties (Model)
STU	Soil Typological Unit
SWIR	Shortwave Infrared
TAGEE	Terrain Analysis in Google Earth Engine
TFO	The Future Okavango (Project)
TIFF	Tagged Image File Format
TPI	Topographic Position Index
TRI	Terrain Ruggedness Index
TWI	Topographic Wetness Index
UHH	(laboratory of the) University of Hamburg's Institute of Soil Science,
UNAM	University of Namibia
UNCBD	United Nations Convention on Biological Diversity
UNCCD	United Nations Convention to Combat Desertification
UNEP	United Nations Environmental Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNFCCC	United Nations Framework Convention on Climate Change
USDA	United States Department of Agriculture
USGS	United States Geological Survey
Vis-NIR	Visible and Near-Infrared Spectroscopy

WCS	Web Coverage Service (for accessing and manipulating raster data)
WFS	Web Feature Service (for accessing and manipulating vector data)
WGS84	World Geodetic System 1984
WMS	Web Map Service (to facilitate visualisations of data)
WISE	World Inventory of Soil Emission Potential
WoSIS	World Soil Information Service (curated by ISRIC)
WSI	(ISRIC-) World Soil Information
WRB	World Reference Base for Soil Resources
XML	Extensible Markup Language

1. INTRODUCTION

1.1 Context

Soils underpin terrestrial ecosystems, agricultural productivity, and sustainable land use. Soil information is fundamental to agronomy and food security, informing land suitability assessments, crop selection, fertiliser management, and crop growth and yield modelling. Beyond agriculture, soil data support a wide range of environmental applications, including climate change assessment, sustainable rangeland management, land degradation and drought monitoring, ecosystem function analysis, environmental impact assessment, and land restoration. Policymakers and society require more than raw measurements; they need interpreted, value-added soil information for planning, resource allocation, risk management, site selection, and infrastructure development (Bispo et al., 2014; Creamer et al., 2021; Omuto et al., 2012).

Global environmental conventions — including the UNFCCC, UNCCD, and UNCBD (United Nations, 1992b, 1992a, 1994) — have renewed international attention on the role of soils in climate regulation, sustainable land management, and ecosystem functioning (Hartemink, 2008; Hartemink & McBratney, 2008). The Revised World Soil Charter (FAO, 2015) emphasises the importance of understanding soil properties and their responses to natural and anthropogenic pressures for achieving sustainability. Reliable soil datasets are increasingly recognised as essential for evaluating soil functions, ecosystem services, and environmental change (Adhikari & Hartemink, 2016; Bouma, 2014; Dominati et al., 2010; Greiner et al., 2017; Rabot et al., 2018; Vogel et al., 2018, 2019). At the same time, advances in geospatial technologies, earth observation, communication systems, and modelling have increased global demand for accessible, harmonised, and spatially referenced soil information. Effective grazing management, bush-control strategies, regenerative land management, and biodiversity conservation all require reliable soil information to guide decisions aimed at improving ecosystem services, water cycling, soil health, and agricultural productivity.

In Namibia, the need for soil information is amplified by the country’s arid to semi-arid climate, fragile ecosystems, and strong dependence on natural resources. Agriculture and nature-based tourism form important pillars of the economy and rural livelihoods, despite soils often being inherently low in fertility and constrained by drought, salinity, shallow depth, or limited water-holding capacity. Effective grazing management, bush-control strategies, regenerative land management, and biodiversity conservation all require reliable soil information to guide decisions aimed at improving ecosystem services, water cycling, soil health, and agricultural productivity.

Namibia is highly vulnerable to climate change and land degradation, both of which pose direct risks to food security and environmental sustainability. The country is a signatory to the three Rio Conventions and has developed a range of sound environmental policies and action plans addressing climate change, biodiversity conservation, and land degradation (Republic of Namibia, 2007, 2011, 2012, 2014c, 2014a, 2014b). As a member of FAO, Namibia is obliged under the Revised World Soil Charter to “develop and maintain a national soil information system and contribute to the development of a global soil information system” (FAO, 2015). These obligations and policy frameworks increase the need for nationally coherent soil information capable of supporting environmental monitoring, land-use planning, and international reporting obligations.

1.2 Problem statement

Despite the strategic importance of soil information, soil science and systematic soil research in Namibia have historically received limited institutional recognition and support. Agricultural research receives less than 4% of the agricultural budget, which itself accounts for only about 5% of the national budget — well below the 10% regional commitment under the Maputo Declaration (African Union Assembly, 2003; Benin & Yu, 2013; Goyal & Nash, 2017). Soil research receives a minute fraction of this already constrained agricultural research funding. As a result, the country has experienced a severe shortage of trained soil scientists, investment in soil information infrastructure, and institutional capacity. The contributions of soils to crop productivity and rangeland restoration are under-researched and soil remains largely absent from national planning policies and plans.

Since the 1950s, several *ad hoc* surveys have been conducted (AOC Technical Services, 1967a, 1967b, 1967c, 1966; Department of Water Affairs, 1987; Du Plessis and Dohse, 1978; Ellis et al., 1980; Lambrechts, 1977; Lambrechts & Schloms, 1993; Louw, 1957; Loxton et al., 1974a, 1974b, 1970; MAWRD, 1995; Murray et al., 1997; Opperman et al., 1982; Page, 1980, 1979, 1978, 1976; Partridge, 1987a, 1987b; Van Woerkom & Louw, 1971). However, much of Namibia's soil information has historically been project-specific, scattered across institutions, consultancy reports, academic theses, unpublished reports, and other forms of grey literature. Many datasets remain in analogue form or obsolete digital formats, while associated field notes, maps, laboratory records, and metadata are vulnerable to deterioration or loss. Differences in sampling designs, analytical protocols, classification systems, spatial coverage, and data structure further complicate integration and reuse.

The gradual loss of institutional memory associated with past survey campaigns poses an additional threat. As researchers retire and institutions change, undocumented knowledge relating to sampling methods, classification decisions, laboratory procedures, and map interpretation risks being permanently lost. Some original field forms, maps, and datasets have already disappeared, limiting the ability to verify or reinterpret legacy observations.

Namibia lacks a complete national soil map, a harmonised and publicly accessible national soil profile database, and soil property maps. The absence of an integrated national soil information system constrains data accessibility, interoperability, long-term preservation, and the practical use of soil information for land assessment, agricultural planning, environmental management, climate adaptation, and national reporting obligations. With no new state-funded soil surveys planned, future progress depends heavily on the rescue, harmonisation, and revitalisation of legacy soil information.

There is, thus, a clear need for spatially coherent and reliable soil information to support agricultural productivity, biodiversity conservation, climate change adaptation, land degradation control, ecosystem service preservation, natural resource governance, and national development planning in Namibia. Meeting this need requires consolidating dispersed legacy soil information into an accessible digital system, completing and harmonising the unpublished preliminary national soil map, filling spatial gaps in soil property knowledge, and aligning the information with both national development priorities and global environmental reporting.

2. AIM AND OBJECTIVES

2.1 Research aim

The overarching aim of this study is to transform Namibia's fragmented legacy soil information into an integrated national soil information system that supports both conventional and digital soil mapping, thereby enabling coherent, accessible, and spatially explicit soil knowledge for national-scale and global applications.

2.2 Specific objectives

To achieve this aim, the study pursues four interrelated objectives:

1. Development of a national soil information framework

To conceptualise, design, and implement an open-access and user-friendly Namibian Soil Information System (NamSIS) that integrates quality-assessed and harmonised soil profile data and supporting resources, compatible with global soil information systems and modelling frameworks.

2. Legacy data recovery and consolidation

To identify, collect, digitise, and secure Namibian soil documents and legacy datasets, ensuring their long-term preservation and accessibility through a structured and searchable archive.

3. Conventional soil mapping

To produce an updated national conventional soil map of Namibia, classified according to the 2015 World Reference Base (WRB) for Soil Resources, providing a coherent pedological framework suitable for national and regional applications.

4. Digital soil modelling

To develop predictive spatial models of selected soil properties using machine learning approaches, generating digital soil maps and associated uncertainty estimates that complement and extend the conventional mapping framework.

2.3 Study approach and hypotheses

The study is framed around the pragmatic integration of complementary epistemological approaches to soil knowledge, transforming multiple sources of evidence — empirical observations, expert interpretation, and data-driven inference — into a unified information system.

Because the study follows an information integration paradigm rather than a statistically testable experimental design, it is guided by a set of qualitative conceptual and operational hypotheses.

Overarching hypothesis

Integrating Namibia's fragmented legacy soil information with global and locally derived

environmental covariates — through harmonisation, conventional soil mapping, and machine-learning-based digital soil mapping — produces more consistent, accessible, and nationally useful soil information than previously available.

The overarching hypothesis is operationalised through three subsidiary hypotheses:

H₁. Data integration hypothesis

Harmonising and integrating fragmented legacy soil data into a structured national soil information system (NamSIS) improve the accessibility, interoperability, long-term preservation, and utility of Namibia's soil information.

H₂. Pedogenic-environmental relationship hypothesis

Spatial variability in soil properties across Namibia can be predicted from climate, terrain, vegetation, and parent-material covariates because dryland pedogenesis is strongly conditioned by environmental gradients.

H₃. Conventional-digital soil mapping convergence hypothesis

Combining expert-driven conventional mapping with machine-learning approaches produces more internally consistent and interpretable national soil information than either approach alone.

These hypotheses guide the design of the three major research components: the conceptualisation of a national soil information system with integrated legacy data, the production of a national soil map, and the modelling of soil properties using digital approaches.

2.4 Dissertation structure

This dissertation is organised into eight chapters:

Chapter 1 introduces the study by outlining the global and national importance of soil information and identifying the research problem.

Chapter 2 defines the research aim and objectives and poses a set of operational hypotheses.

Chapter 3 provides an overview of the national study area and presents the context and literature review. It provides overviews of typical dryland pedogenic processes, global and continental soil information and mapping initiatives, and legacy soil data. It also synthesises the theoretical foundations and methodological considerations relevant to soil information systems, and conventional and digital soil mapping. It concludes with a discussion on current debates and knowledge gaps.

Chapter 4 describes the methods used in the study. It follows the research objectives sequentially, detailing the conceptualisation and design of the Namibian Soil Information System, the rescue, renewal and harmonisation of legacy soil data, the compilation of the national soil classes map, and the development of predictive spatial models of selected soil properties using machine learning techniques.

Chapter 5 presents the results of the study: a functional Namibian soil information system, including the consolidated digital archive of legacy documents, maps and raw data, the national profile database, the national soil map, predictive digital soil property maps with uncertainty

estimates, modelling scripts, and metadata.

Chapter 6 provides a discussion of the results in relation to the research objectives, methodological considerations, limitations, and wider national and global contexts.

Chapter 7 draws conclusions about the primary outcomes, support for the hypotheses, and scientific significance of the study.

Chapter 8 provides recommendations for future applications, maintenance, and continued development of the information system, and addresses cross-cutting issues.

The *References* list all sources cited in the dissertation and the *Appendices* contain supplementary materials.

3. CONTEXT AND LITERATURE REVIEW

This chapter briefly summarises environmental conditions and land use in Namibia, as well as the main dryland pedogenesis processes and soil types. It then reviews the conceptual foundations for this research: global soil information initiatives, soil information systems, legacy data revitalisation, conventional and digital soil mapping approaches. Together, these contextualise the development of a Namibian Soil Information System, including the compilation of a harmonised national soil profile database, production of a national soil map, and generation of machine-learning-based digital soil property predictions.

3.1 Environmental conditions and land use in the Republic of Namibia

Namibia is situated on the Atlantic coast of southern Africa, straddling the Tropic of Capricorn between approximately 11°45'–21°00' E and 16°56'–28°58' S, with the narrow Caprivi Strip extending eastwards to 25°15' E. The country covers roughly 823 450 km², with elevations ranging from sea level along the coast to 2 573 m at Königstein Peak in the Brandberg massif (Figure 1).

Four broad physiographic regions define its landscape (Figure 2): The Namib Desert borders the Atlantic Ocean, runs the length of Namibia's coastal plain and comprises dune fields, extensive gravel plains, and scattered inselbergs. The rugged Great Escarpment rises steeply from the coastal plain to the highlands of the interior, the Central Plateaus (~900–2 500 m). To the north, northeast and southeast, the land slopes down to the Kalahari Basin (Atlas of Namibia Team, 2022). The main landscapes are displayed in Figure 3 and broad lithological classes in Figure 4

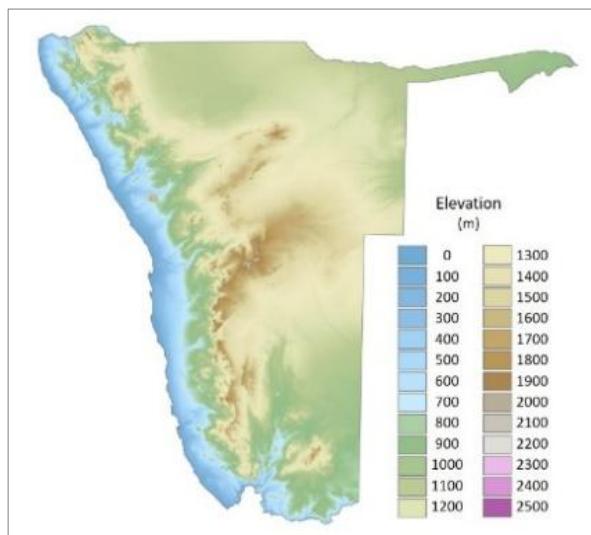


Figure 1. Elevation of Namibia above mean sea level.

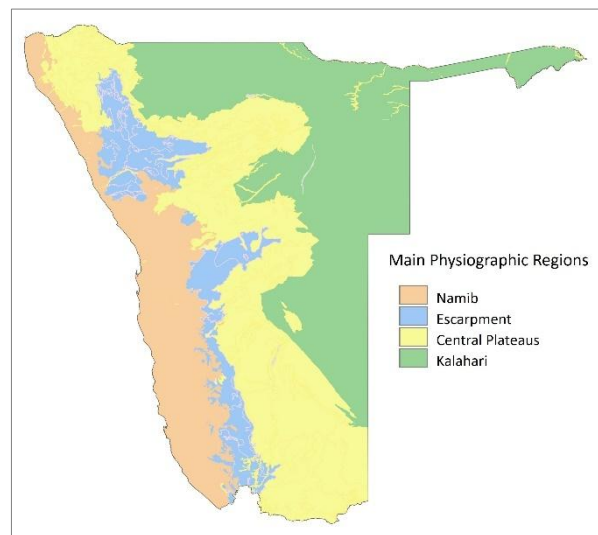


Figure 2. Main physiographic regions of Namibia.

The present geological surface is a mosaic reflecting the country's long geological history. Precambrian crystalline basement rocks, consisting mainly of granites, gneisses, and other high-grade metamorphic rocks belonging to the Congo and Kalahari cratons, crop out in northern and western Namibia. The central highlands are largely composed of metamorphosed marine sediments such as mica-rich schists, quartzites, and marbles that were originally deposited in an ocean between the two cratons and subsequently metamorphosed during the Neoproterozoic

Damara Orogeny associated with the Pan-African assembly of Gondwana. Granitic intrusions emplaced during this period now form many of the prominent inselbergs in central Namibia. The Owambo and Nama foreland basins in northern and southern Namibia, respectively, contain limestones, dolomites, and shales deposited on shallow marine shelves. Subsequent erosion produced a broad Gondwana erosion surface. During the late Palaeozoic and early Mesozoic, continental plate movements led to successive phases of glaciation, deposition of glacial, lacustrine and fluvial sediments (tillites, shales, mudstones), and development of large deserts with widespread aeolian sand deposition. These deposits collectively form the Karoo Supergroup. Jurassic-Cretaceous magmatism associated with the break-up of Gondwana produced extensive flood basalts, latites, dolerite sills, and intrusive igneous complexes. Post-break-up isostatic up-warp of the continental margins accelerated erosion and accumulation of vast volumes of sediment within the Kalahari Basin, forming the Kalahari Group. More recently, sediment transport from the interior of southern Africa and coastal longshore drift led to the development of the Namib Sand Sea (Atlas of Namibia Team, 2022; Haddon & McCarthy, 2005; Miller, 2008).

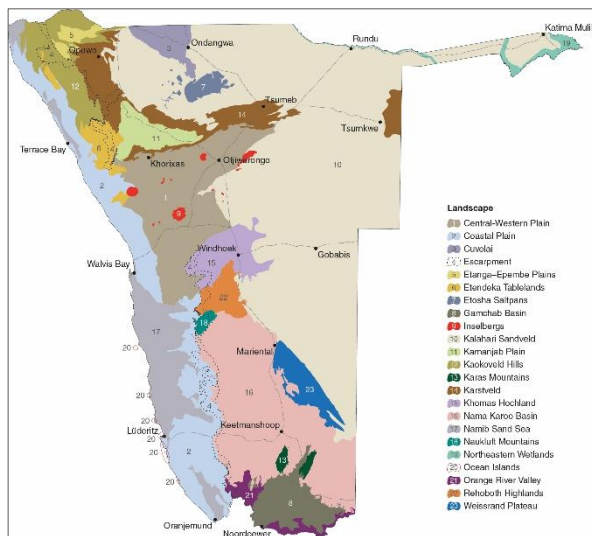


Figure 3. Landscapes of Namibia (Atlas of Namibia Team, 2022)

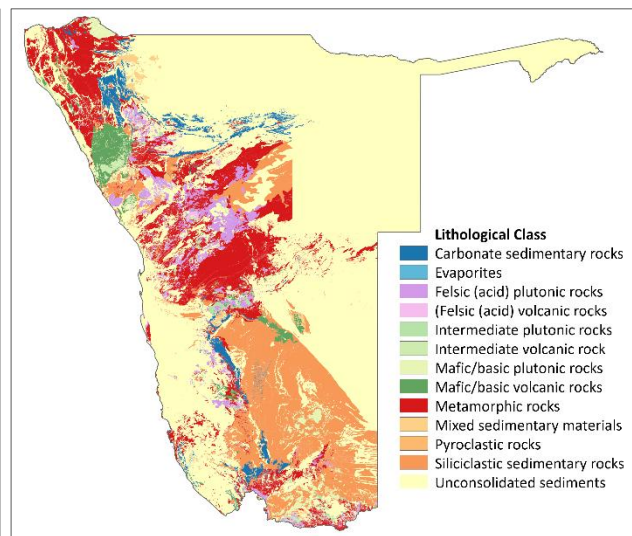


Figure 4. Broad lithological classes of Namibia (derived from Geological Survey of Namibia, n.d.) (mapping incomplete in northeast)

Namibia is the most arid sub-Saharan African country, characterised by extreme spatial and temporal rainfall variability. Median annual rainfall ranges from less than 50 mm in the Namib Desert to around 650 mm in the far northeast (Figure 5), with interannual variability increasing from northeast to southwest. While the rainy season can stretch from October to April, peak rainfall typically occurs from December to March, with June to August is the coolest and driest period (Atlas of Namibia Team, 2022).

The Global Aridity Index (GAI) classifies most of Namibia as arid (GAI 0.04–0.20), with the Namib Desert falling into the hyper-arid category (GAI <0.04), and the northeast considered semi-arid (GAI 0.20–0.28) (Figure 6) (Atlas of Namibia Team, 2022). Namibia's aridity stems from its position in the subtropical high-pressure zone and the effect of the cold, northward-flowing Benguela Current, which exhibits strong upwelling along the Namibian coast and suppresses evaporation. Apart from overnight and early-morning coastal fogs, there is no appreciable influx of moisture over the country from the west.

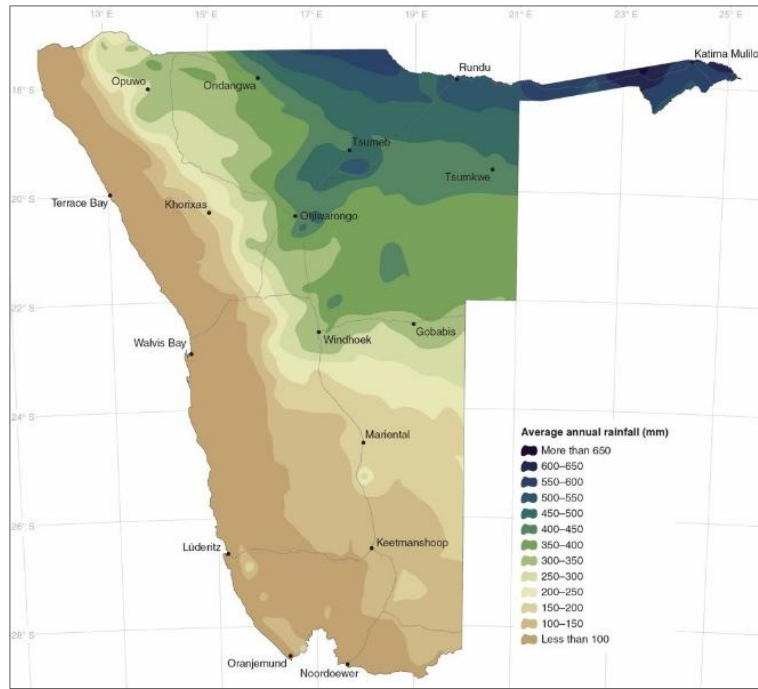


Figure 5. Mean annual rainfall of Namibia (Atlas of Namibia Team, 2022).

The South Atlantic and Botswana Anticyclones dominate in winter, inhibiting rainfall. Only the far south receives occasional winter rain from temperate zone low-pressure systems and cold fronts sweeping from west to east across southern Africa. In summer, the anticyclones weaken and the low-pressure Intertropical Convergence Zone (ITCZ) shifts southwards, following the sun and carrying warm, moist air from equatorial regions. The interplay between the ITCZ and the anticyclones determines when and how much rain Namibia will receive in a rainy season (Atlas of Namibia Team, 2022).

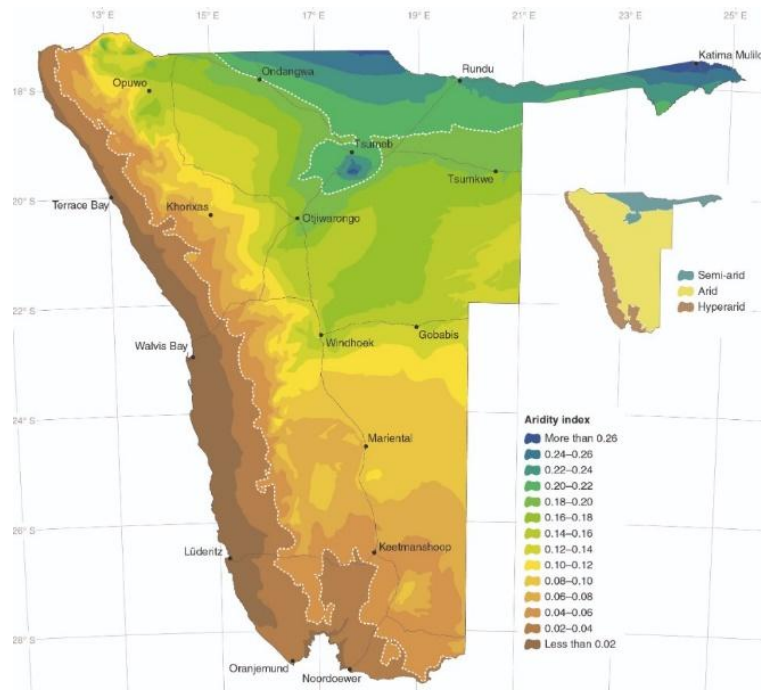


Figure 6. Aridity index of Namibia (Atlas of Namibia Team, 2022).

Rainfall is modulated by atmospheric teleconnections, namely the Indian Ocean Dipole,

Subtropical Indian Ocean Dipole, South Atlantic Oscillation Dipole, and El Niño–Southern Oscillation, with El Niño generally associated with drought and La Niña with above-average rainfall (Du Pisani, 2001).

Potential evaporation is among the highest in Africa, ranging from ~1 800 mm/year along the fog-influenced coast to over 2 800 mm/year in the south, and peak evaporation occurring during the hottest months with low cloud cover (October to December) (Atlas of Namibia Team, 2022). Only 2% of annual rainfall becomes surface runoff, 1% recharges aquifers, 14% supports plant growth, and ~83% simply evaporates (IWRMP Joint Venture Namibia, 2010).

Vegetation and ecosystems reflect these climatic and topographic gradients. Namibia falls within the Deserts and Semi-Deserts (T5) and Savannas and Grasslands (T4) biomes (Keith et al., 2022; Strohbach et al., 2025). The more humid northeast supports tree-dominated savannas and open woodlands with low canopy heights and a well-developed herbaceous layer, while shrubland and dwarf shrub savannas are typical of the drier western, central and southern areas (Figure 7). The Namib Desert’s gravel plains and inselbergs host resilient species adapted to extreme aridity. Alluvial aquifers associated with ephemeral rivers create linear oases across the desert from the highlands and escarpment to the coast. The southwest is a centre of high endemism with succulents and dwarf shrubs associated with small but seasonally important winter rainfall.

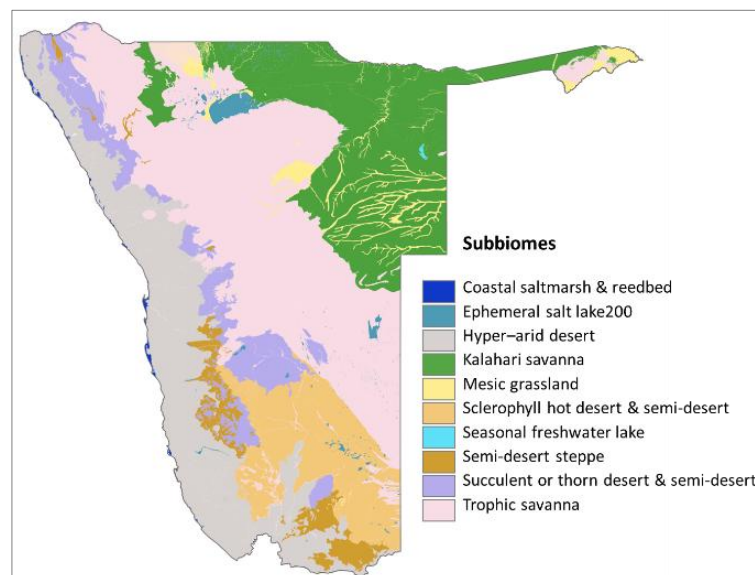


Figure 7. Sub-biomes of Namibia (based on Strohbach et al., 2025).

Agriculture is spatially dominated by extensive grazing, predominantly ranching and sedentary pastoralism, with limited transhumance (Figure 8). Climatologically, only ~2% of Namibia is suitable for rainfed crop production, and an additional ~6% moderately suitable (De Pauw, 1996). There are few large, commercial crop enterprises due to water scarcity, climatic limitations and the land tenure system, thus cropland is mostly mosaics of small, low-input subsistence crop fields interspersed with rangeland in the north and northeast of the country. The far west and southwest are preserved as natural habitats with low-impact landscape and wildlife tourism. According to the Copernicus Land Cover data (Buchhorn et al., 2020a; Buchhorn et al. 2020b), the main land cover types are shrubs (51.2%), bare/sparse vegetation (27.6%), herbaceous vegetation (14.2%), and open and closed forests (5.4%), while only 0.85% of land is cultivated, permanent water bodies occupy 0.02%, and infrastructure 0.03%.

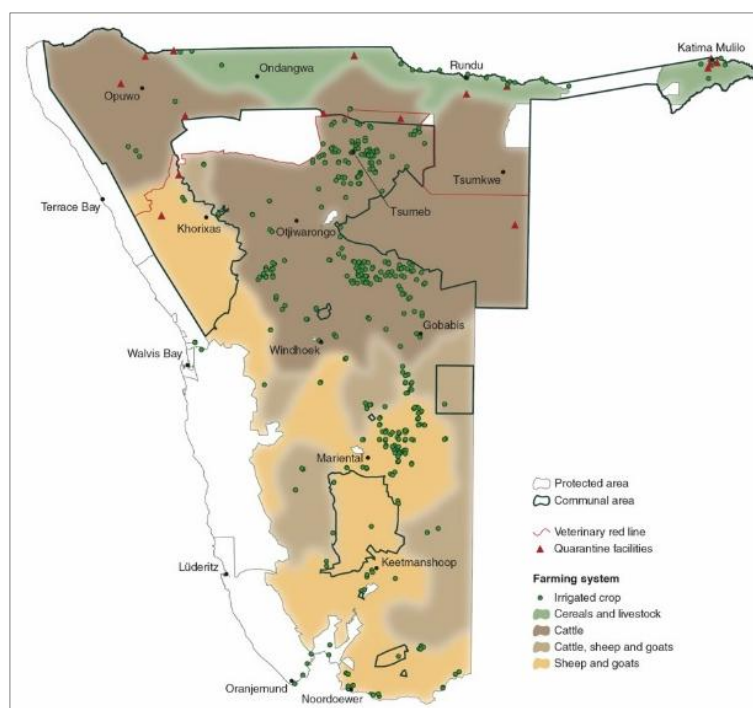


Figure 8. Major farming systems of Namibia (Atlas of Namibia Team, 2022).

Land degradation is a major concern: approximately 56% of the country (~460 000 km²) is affected by bush thickening, a form of rangeland degradation where unpalatable, mostly thorny, indigenous shrubs displace grasses and other desirable forbs and fodder bushes, causing a decrease in productive capacity, ecosystem services, soil health, and biodiversity (MEFT, 2024). Most rangelands also exhibit other forms of degradation: loss of diversity of desirable perennial grasses, forbs and fodder bushes, decreased biomass production, lower soil organic matter, biotic depletion, desiccation, surface crusting, and erosion (Republic of Namibia, 2012, 2014c). These environmental pressures — combined with climatic variability and inherently low soil fertility — make the availability of reliable soil information essential for sustainable land management, climate adaptation, and agricultural planning.

3.2 Major dryland pedogenic processes and the influence of lithology

The terminology used below follows the World Reference Base for Soil Resources (WRB), in which diagnostic horizons, properties, and materials provide a standard basis for soil classification and mapping (IUSS Working Group WRB, 2015).

Physical disintegration of rock and sediment is important in drylands because large diurnal temperature ranges, wetting and drying, abrasion by wind-blown particles, and salt crystallisation fracture and weaken mineral materials (Goudie, 2013; Schaetzl & Thompson, 2015). Freeze-thaw weathering contributes locally where temperatures occasionally drop below freezing point, but it is not a dominant process throughout Namibia. Chemical weathering is not absent in drylands, but is spatially and temporally discontinuous because hydrolysis, dissolution, oxidation, hydration, and carbonation depend on the availability and residence time of water in the soil profile (Brady & Weil, 2016; Schaetzl & Thompson, 2015). Even short wet periods can produce significant chemical transformations, especially in finer-textured materials, zones of water accumulation, and areas influenced by shallow groundwater (Royal, 2023; Verheye, 2022).

Soil development must be interpreted in relation to local climate, parent material, landform, drainage conditions, vegetation, and the time for which a surface has remained geomorphically stable (Schaetzl & Thompson, 2015; Goudie & Viles, 2015). Pedogenesis in arid and semi-arid environments is influenced strongly by the balance between water inputs and potential evapotranspiration (Lambrechts & MacVicar, 2004; Royal, 2023; Verheye, 2022). In Namibia, low and highly variable rainfall, high evaporative rates, high temperatures, sparse vegetation cover, and frequent water deficits restrict the rate and intensity of soil-forming processes (Goudie & Viles, 2015; Goudie, 2013), but their influence varies across Namibia's climatic, geomorphic, and geologic gradients. The hyper-arid Namib, seasonally influenced interior, Kalahari Sandveld, north-central pans, and relatively wetter north-eastern regions do not experience identical soil-moisture regimes or pedogenic pathways.

Limited net downward percolation constrains the removal of soluble substances from profiles. Thus, many dryland soils retain weatherable primary minerals and show weak horizon differentiation relative to soils formed under persistently humid conditions (Brady & Weil, 2016; Schaetzl & Thompson, 2015). Although leaching is restricted, episodic infiltration does dissolve and translocate carbonates, gypsum, salts, clay, iron, and organic compounds over short distances within the profile. The depth and form of accumulation depend on water movement, the solubility and supply of the respective constituents, soil texture and structure, the presence of restrictive layers, groundwater influence, and the duration of landscape stability (Driessen et al., 2001; Fey, 2010; Goudie, 2013). Repeated cycles of dissolution and reprecipitation of calcium carbonate lead to the development of calcic horizons in many arid and semi-arid soils, frequently accompanied by decalcification of surface layers (Ellis, 1988, 1990; Ellis & Lambrechts, 1994; Fey et al., 2010; IUSS Working Group WRB, 2015). The sources of carbonate may be calcareous parent material, dust inputs, groundwater, or redistribution within the soil profile (Driessen et al., 2001; Fey et al., 2010; Lambrechts & MacVicar, 2004; Schaetzl & Thompson, 2015). A petrocalcic horizon forms where carbonate accumulation becomes sufficiently indurated. This process, calcretisation, is favoured by a prolonged period of relative geomorphic stability but can be interrupted or modified by erosion, sediment deposition, dissolution, or changes in climate and hydrology (Fey et al., 2010; Goudie, 2013; Goudie & Viles, 2015).

Secondary gypsum accumulation can occur where calcium and sulfate are available and where evaporation exceeds effective leaching, producing a gypsic horizon or more strongly cemented petrogypsic horizon (IUSS Working Group WRB, 2015). These features are especially relevant to Namibia's most arid environments, namely parts of the Namib Desert (Heine & Walter, 1996; Miller, 2008). Similarly, silica accumulation may produce duric or petroduric horizons under long-term landscape stability and specific geochemical conditions — high alkalinity, sufficient silica supply and water movement (Ellis, 1988, 1990, 2002; Ellis & Lambrechts, 1994; Fey et al., 2010; Francis et al., 2007; Francis, 2008; IUSS Working Group WRB, 2015).

Salt accumulation and sodification occur most strongly where water high in dissolved salts is concentrated by evaporation or where shallow saline or sodic groundwater influences the profile (Lambrechts & MacVicar, 2004; Royal, 2023; Verheye, 2022). Both salinisation (accumulation of readily soluble salts) and sodification (the dominance of exchangeable sodium on the soil exchange complex) adversely affect soil structure, infiltration, and plant-water availability (Brady & Weil, 2016; Driessen et al., 2001; Fey et al., 2010). In Namibia, these processes are particularly noticeable in closed depressions and poorly drained lowlands, and in coastal and interior salt pans such as parts of the Cuvelai-Etosa system. They are landscape-specific outcomes of hydrology and solute redistribution rather than characteristics of all dryland soils.

On steep or rocky terrain, erosion may remove material faster than soil horizons can develop, producing shallow and/or stony soils (Lambrechts & MacVicar, 2004; Royal, 2023; Verheye, 2022). Conversely, profile development can be interrupted by deposition on footslopes, alluvial plains, and aeolian surfaces, which rejuvenates soil parent materials (Goudie & Viles, 2015; Schaetzl & Thompson, 2015). While wind erosion selectively removes fine particles from exposed surfaces in some positions, dust and sand deposition contributes mineral materials, nutrients, and carbonate elsewhere (Goudie, 2013). Surface characteristics of dryland soils commonly include desert pavements, vesicular crusts, and biological soil crusts — mainly composed of cyanobacteria, lichens, and mosses — which influence infiltration, erosion, and nutrient cycling (Fey et al., 2010). All these processes contribute to local contrasts in texture, depth, surface stoniness, water-holding capacity, and nutrient status, shaping the diversity of Namibian soils.

Arid and semi-arid conditions limit biomass production, which results in low organic-matter inputs in dryland soils, while concurrent warm conditions accelerate mineralisation (Brady & Weil, 2016; Fey et al., 2010). Organic carbon accumulation is therefore generally modest in well-drained arid landscapes, though local exceptions occur where vegetation is more productive, sediment is deposited, groundwater is accessible or flooding and poor drainage reduce decomposition. Similarly, eluviation and illuviation of clay, iron, and organic matter are generally less extensive where rainfall is insufficient to sustain prolonged percolation, yet they may be locally important in wetter regions, fine-textured soils, and sites receiving concentrated water flow (Fey et al., 2010; Schaetzl & Thompson, 2015).

Although much of Namibia is characterised by water limitation and evaporative concentration, hydromorphic processes do occur in wetlands, pans, drainage channels, seasonally inundated floodplains, and areas affected by shallow groundwater. Periodic or prolonged saturation can lead to reducing conditions and associated redoximorphic features, creating sharp pedological contrasts with adjacent freely drained soils (IUSS Working Group WRB, 2015; Fey, 2010; Goudie & Viles, 2015).

Lithology influences Namibian soils principally through the grain size, mineral weatherability and chemical composition of the parent material. Along the broad siliceous-to-mafic compositional gradient, quartz-rich aeolian sands, quartzites and quartz sandstones contain few weatherable minerals, while granites and rhyolites are also relatively rich in silica and poor in Ca, Mg and P (Gray et al., 2016). These materials tend to produce coarse textured, weakly buffered soils with low clay content, cation exchange capacity, water holding capacity and inherent fertility, particularly Arenosols, Regosols and, where erosion restricts soil depth, Leptosols. Mafic rocks such as basalt, dolerite and gabbro, together with mafic metamorphic rocks such as amphibolite, contain more readily weatherable ferromagnesian minerals and Ca-rich plagioclase and generally supply more base cations, Fe and P. Their weathering therefore tends to produce finer-textured, Fe-rich and more base-saturated materials, commonly containing smectitic or other 2:1 clay minerals. Depending on drainage, relief and profile age, these may support Cambisols, Luvisols or shrink-swell soils with vertic properties. Gray et al. (2016) found lithology to be the most influential covariate for pH, clay content, CEC, sum of bases, total P and soil organic carbon. Their chemical sequence places quartz sand and quartzite at the highly siliceous end; granite, rhyolite and sandstone in siliceous classes; shale and schist in an intermediate class; and basalt, dolerite, gabbro and amphibolite in the mafic class. Increasing silica was associated with declining clay, bases, CEC, P and soil organic carbon. Fine-grained shale, mudstone and mica-rich schist also produce clay-rich soils, but their properties vary with the abundance of illite, mica, chlorite and expandable clays. Navarro-Hasse et al. (2023) deduced that metamorphic rocks such as schists

and gneisses should not be interpreted as a uniform parent material class because their clay products depend strongly on the minerals of the original protolith. Limestone, dolostone and marble contribute Ca and Mg. Their dissolution can leave shallow, stony residual soils on exposed uplands, while carbonate redistributed into more stable landscape positions produces calcic or petrocalcic horizons characteristic of Calcisols.

The relationships between geology and soil are not always direct. Aeolian, fluvial, colluvial, and lacustrine deposits conceal the underlying bedrock across much of Namibia, while aridity, landform, drainage and landscape stability strongly modify the properties inherited from the geological substrate (Atlas of Namibia Team, 2022; Driessen et al., 2001; Fey, 2010; Miller, 2008). A recent Namibian study of five profiles in the Otavi Mountains Chromic Cambisols, Haplic and Vertic Phaeozems and a Calcic Kastanozem within a small area. The sharp soil contrasts were not simply controlled by the underlying limestone and dolomite, but rather by transported parent materials, inherited and newly formed clay minerals, subtle relief differences and water stagnation (Zádorová et al., 2025). Geological maps therefore provide an important predictor of soil properties, but must be interpreted together with regolith transport, landform, drainage and aridity (Gray et al., 2016; Zádorová et al., 2025; Eckardt et al., 2001).

Namibia's soil diversity is the result of interacting climatic, lithological, geomorphic, hydrological, and biological conditions rather than of aridity alone. These controls generate spatial variation in soil classes and properties, providing the pedological basis for the conventional interpretation of soil landscapes and for the use of environmental covariates in predictive modelling.

3.3 Main soil types of Namibia

The following overview of Namibia's principal soil types synthesises information from the 1998–2000 National Soil Survey (ICC & MAWRD, 2000) and the National Agricultural Research Information System (NARIS) (MAWRD-AEZ, 2009), supplemented by diagnostic information from the WRB (IUSS Working Group WRB, 2015), the FAO Lecture Notes on the Major Soils of the World (Driessen et al, 2001), and other cited studies.

Arenosols are deep sands or loamy sands, formed in aeolian, lacustrine, marine and littoral deposits, or through in situ weathering of quartz-rich sediments and rocks. They are typically unconsolidated with weak, single-grain structure and limited horizon development. Bulk densities commonly range from 1.5–1.7 g.cm⁻³, with total porosities of about 36–46%. Their high macroporosity results in good aeration and rapid drainage, but also in low water retention. Organic carbon contents are generally very low, often less than 0.5% in topsoil and less than 0.2% in subsoil in Namibia. Most soluble constituents are leached beyond the rooting zone of shallow-rooted plants, leading to frequent decalcification of the surface layer. Because quartz is chemically rather inert and these soils contain few clay or organic matter colloids, cation exchange capacity and buffering capacity are low, and inherent fertility is poor. Reddish and yellowish colours are imparted by iron oxide coatings (hematite, goethite) on sand grains (Driessen et al., 2001; Fey, 2010; IUSS Working Group WRB, 2015). Although total clay contents are low, the Namib Arenosols contain relatively higher proportions of smectitic and vermiculitic clays than the more weathered Kalahari Arenosols, where higher leaching favours kaolinitic clay minerals (Heine & Völkel, 2010). The beach and dune sands of the northern Namib contain up to 80% quartz, along with feldspars, minor micas, ferromagnesian minerals and heavy minerals such as zircon, garnet,

tourmaline, ilmenite, magnetite and rutile (Garzanti et al., 2014).

Calcisols occur throughout Namibia, reflecting the importance of secondary calcium carbonate accumulation as a soil-forming process under arid and semi-arid conditions. Calcisols typically have a thin, pale brown surface horizon that may show granular or crumb structure, vesicular or biological crusts, and partial decalcification due to episodic leaching. Repeated cycles of wetting and drying lead to the downward transport and concentration of CaCO_3 in a subsurface calcic horizon, expressed as dispersed calcite, soft or hard nodules, veins, pseudomycelia or coatings. These horizons are commonly white, pinkish or greyish, and often have reduced porosity because soil pores are partially or completely filled with carbonates. With time, calcic horizons can become strongly cemented into petrocalcic horizons (calcrete), which may be nodular, lamellar or massive. Calcisols have a high base saturation and neutral to strongly alkaline pH, reaching 8.0–8.5 where free carbonates are abundant (Driessen et al., 2001; Fey, 2010; IUSS Working Group WRB, 2015).

Cambisols occur widely in weakly developed, medium- to fine-textured colluvial, alluvial, and aeolian deposits. They are associated either with relatively recent sediment deposition or with environments where aridity inhibits pedogenic development. Cambisols show early signs of parent material weathering and horizon differentiation, such as the onset of structure formation, colour differentiation, and clay or carbonate redistribution. The cambic horizon of a Cambisol formed in situ typically has a slightly higher clay and/or sesquioxide content than underlying material and is often partially decalcified (Driessen et al., 2001; IUSS Working Group WRB, 2015).

The formation of *Durisols* is associated with prolonged geomorphic stability, alkaline conditions, and the accumulation and cementation of secondary silica. In southern Namibia, their development seems to also be associated with winter rainfall and termite-mediated processes (Ellis, 1988, 1990, 2002; Ellis & Lambrechts, 1994; Fey, 2010). These Durisols are generally reddish-brown to very pale brown, shallow to moderately deep, moderate to coarse-textured (sandy clay loam to sand), and have relatively low water-holding capacity. Duric/petroduric layers may be only a few centimetres thick or up to several metres thick, with massive or platy structure that impedes deep percolation. Roots often form mats above these horizons or grow laterally along fractures and old burrows. Durisols usually have high base saturation, relatively high exchangeable sodium and extractable boron (sometimes exceeding toxic thresholds), low organic carbon, and low extractable iron (Francis, 2008; Francis et al., 2007). Bulk density can be high, particularly in petroduric layers. Duric and petroduric horizons commonly grade laterally into gypsic, petrogypsic, calcic and petrocalcic horizons, and durinodes are often coated with carbonates (Driessen et al., 2001; Ellis, 1988, 1990, 2002; Fey, 2010; IUSS Working Group WRB, 2015).

Fluvisols develop in unconsolidated fluvial, lacustrine, and marine sediments. They display stratification in texture, colour, organic matter content and coarse fragments that reflect varying energy conditions during deposition. Some Fluvisols have takyr properties such as compact, fine-textured crusts with prominent polygonal cracks developed over underlying material with a platy or massive structure (Driessen et al., 2001; IUSS Working Group WRB, 2015).

Gypsisols are typical soils of the Namibian coastal gravel plains, usually associated with the alteration of Tertiary and Pleistocene calcretes and crusts (Heine & Walter, 1996; Miller, 2008). Their formation is linked to the Benguela Current upwelling system, onshore transport of atmospheric sulfates of marine biogenic origin by prevailing winds, and coastal fog providing the moisture to carry surface deposits into the soil (Eckhardt et al., 2001; Goudie & Viles, 2015; Miller, 2008). The best-developed Gypsisols are found downwind of the strongest upwelling cells, particularly in the Central Namib. Their expression diminishes further north, due to lower fog

incidence and increased wind intensity that disrupts the stability of the coastal plain pediments (Miller, 2008). Generally, the upper 0.5–1 m contains well-formed crystals, while the subsurface layer, up to 5 m thick and up to 10 m deep, consists of massive crusts of impure gypsum cementing poorly sorted sediments. The surface may be covered by a thin layer of relatively gypsum-free sand (Miller, 2008). Large polygonal patterns, up to 20 m in diameter, are on occasion observed on the soil surface.

Leptosols are shallow and/or extremely stony soils, occurring where active erosion outpaces soil formation. Differential weathering, illuviation and biotic disturbance cause large spatial heterogeneity in these soils. The limited rooting depth and presence of stones limit their water-holding capacity. Leptosols frequently dominate crests and steep slopes while Regosols occupy concave landscape positions and footslopes (Driessen et al., 2001; Fey, 2010; IUSS Working Group WRB, 2015).

Regosols are minimally developed “other soils” that do not satisfy the diagnostic criteria for other reference soil groups. They are commonly found in unconsolidated, moderately fine-textured materials where soil formation is inhibited by aridity, frequent erosion or continuous deposition. Their properties are largely inherited from their parent materials, with poor horizon differentiation (Driessen et al., 2001; IUSS Working Group WRB, 2015).

Luvisols have an argic horizon of high-activity clays and high base saturation, thus they are moderately to highly fertile. Though their physical properties can vary considerably depending on clay content and structure, they are generally considered to be good agricultural soils (Driessen et al., 2001; IUSS Working Group WRB, 2015).

Solonchaks are common in coastal salt pans, desert playas, endorheic depressions (e.g. Etosha Pan and parts of the Cuvelai–Etosha system), and other poorly drained environments where evaporation exceeds precipitation, with saline groundwater or accumulation of surface inflow. Their salinity inhibits plant growth due to osmotic stress and, in some cases, ion toxicity. Solonchaks often exhibit gleyic features at depth due to prolonged waterlogging (Driessen et al., 2001; IUSS Working Group WRB, 2015).

Solonetz are sodic soils found in clay-rich coastal pan sediments and in the Cuvelai–Etosha drainage system. They have a dark, dense, strongly structured, clay-rich natric horizon high in exchangeable sodium, and sometimes magnesium. The surface of the coarse columnar or prismatic peds are frequently coated by thick, dark clay and organic matter cutans, while their characteristic rounded tops are covered by bleached, powdery silt or fine sand. Natric horizons are hard when dry, dispersive when wet, and have poor hydraulic conductivity, making these soils difficult to manage (Driessen et al., 2001; IUSS Working Group WRB, 2015). They are often associated with Solonchaks and underlying calcic or gypsic horizons.

Vertisols are clay-rich, shrink-swell soils dominated by smectitic clays, with deep cracks in dry seasons, and slickensides and wedge-shaped peds in the subsoil (Driessen et al., 2001; Fey, 2010; IUSS Working Group WRB, 2015). They can be naturally fertile but are difficult to cultivate due to the extreme changes in consistency.

Isolated occurrences of *Kastanozems* have been reported near Otjiwarongo (Eitel et al., 2002; Eitel & Eberle, 2001), occurring in landscape positions between Calcisols (upslope) and Vertisols (downslope), where they developed in fine-grained host sediments. These are interpreted as palaeosols formed under shortgrass continental steppe conditions that no longer prevail. Haplic and Vertic *Phaeozems* were recorded in the Otavi Mountains (Zádorová et al., 2025). They are

probably relict soils formed under wet steppe conditions.

3.4 Global and African soil information and mapping initiatives

International efforts to standardise soil information began with the *FAO–UNESCO Soil Map of the World* (FAO/UNESCO, 1971, 1974; FAO/UNESCO/ISRIC, 1990), which provided the first globally consistent depiction of soil classes. Compiled from assorted national inputs and expert interpretation, the map established a common pedological framework that enabled global comparison. It was later digitised as the *Digital Soil Map of the World (DSMW)*, improving accessibility but retaining the coarse scale and generalised nature of the original product (FAO, 1995, 2003). To strengthen the linkage between soils and landscape attributes, the *Soil and Terrain (SOTER)* concept was developed as a standardised approach for combining soil and terrain information within physiographic units (van Engelen & Wen, 1995). SOTER-based databases were produced at national and (sub)continental scales, including *NamSOTER* for Namibia (Coetzee, 2001) and *SOTERSAF* for southern Africa (FAO/ISRIC, 2003). These products advanced soil data harmonisation and comparability and formed a foundation for subsequent global compilations. Notably, SOTER datasets, together with FAO soil maps and other regional sources, formed the basis of the *Harmonized World Soil Database (HWSD)*, which provides global gridded estimates of soil properties at approximately 1 km resolution (FAO & IIASA, 2023; FAO/IIASA/ISRIC/ISSCAS/JRC, 2012). Although often used as a digital soil input layer, HWSD represents an intermediate product that translated categorical mapping units and profile-based attributes into a gridded format, rather than a fully model-driven digital soil mapping output. At the continental scale, the *Soil Atlas of Africa* (Jones et al., 2013) synthesised soil types and properties across Africa, presenting an important reference for education, planning, and broad-scale assessment.

Parallel to map-based products, large initiatives were undertaken to compile and harmonise soil profile observations. The *World Inventory of Soil Emission Potentials (WISE)* database, developed by the International Soil Resources and Information Centre (ISRIC), assembled globally distributed soil profile data and standardised analytical properties to address gaps in map-based datasets (Batjes, 2002, 2004). At the continental scale, the *Africa Soil Profiles Database* consolidated available African data, improving documentation and harmonisation of legacy observations (Leenaars, 2012; Leenaars et al., 2014). These profile databases substantially enhanced the observational and analytical basis for global and continental soil assessment, but accessibility, update mechanisms, and integration with spatial products remained challenging.

Soil information systems (SIS) emerged as integrated infrastructures that combine soil profile data, metadata, harmonisation procedures, access tools, and update mechanisms. ISRIC's *World Soil Information Service (WoSIS)* is a major global SIS, integrating data from WISE, the Africa Soil Profiles Database, and numerous national and project-specific datasets into a quality-assessed, standardised, and continuously updated repository (Batjes et al., 2017, 2020, 2024; Batjes & Ribeiro, 2019; Calisto et al., 2023; Carvalho Ribeiro et al., 2020). WoSIS currently hosts approximately 244 000 standardised soil profiles globally, including a growing number from Namibia, as reported through its dynamic dashboard (ISRIC–WSI, n.d.). At the global governance level, the *Global Soil Information System (GloSIS)* aims to interconnect national SIS platforms into a decentralised global infrastructure coordinated through the FAO Global Soil Partnership and the International Network of Soil Information Institutions (INSII) (De Sousa et al., 2021; FAO,

2020b, 2020a).

Recent African initiatives have focused on addressing the persistent scarcity and uneven distribution of soil observations across the continent. The Horizon 2020 *Soils4Africa* project generated standardised soil profile and analytical data through coordinated field sampling campaigns at more than 15 000 locations across Africa (European Commission, 2020, 2024). These data, together with harmonised legacy datasets, are curated and disseminated through the *Africa Soil Information System (AfricaSIS)*, a continental SIS designed to improve accessibility, interoperability, and long-term stewardship of African soil information (ISRIC–WSI, 2023). AfricaSIS is intended to contribute to the *African Union Soil Observatory (AUSO)*, which seeks to embed soil information management within a continental governance structure that supports agricultural development, land management, and environmental monitoring (Abugri et al., 2025; African Union Commission, n.d.).

The advent of digital soil mapping (DSM) marked a methodological shift from map-based extrapolation toward data-driven, spatially explicit prediction of soil properties. Early global DSM products include *SoilGrids*, which applied machine learning techniques to harmonised soil profile data and environmental covariates to generate global soil property predictions at 1 km resolution (Hengl et al., 2014). Reported prediction performance varied considerably across regions and soil properties, depending on the availability of observed and measured site and profile data. Subsequent developments improved both resolution and regional relevance. *AfSoilGrids250m*, a continent-wide soil property estimates for Africa — produced under the Globally Integrated Africa Soil Information Service (AfSIS) project and using data from the Africa Soil Profiles Database — demonstrated the potential of DSM for data-scarce environments when legacy data are systematically harmonised (Hengl et al., 2015, 2017a, 2021). The *GlobalSoilMap* initiative emphasised the reuse and harmonisation of existing soil profile data as a prerequisite for national digital soil mapping efforts. It formalised DSM standards by defining global specifications for soil property mapping at approximately 100 m resolution across six standard depth intervals, including conventions for uncertainty estimation and data dissemination (Arrouays et al., 2017b, 2020b, 2020c). Together, these initiatives catalysed growth in national soil datasets and established methodological conventions widely adopted in DSM practice.

The effectiveness of continental and global soil information systems and DSM products is affected by the reliability, coherence, and compatibility of contributing national soil information systems. Where national datasets are sparse or fragmented, continental products necessarily rely on generalised mapping units, a small number of observations, and assumptions that may not adequately capture local pedological conditions. Namibia’s representation in global and African soil databases has historically been hampered by limited and unorganised profile data and incomplete national mapping. For example, only 65 Namibian profiles were included in the Africa Soil Profiles Database used for the generation of the original *AfSoilGrids250m* maps, reinforcing the need for a comprehensive national Namibian soil information system.

3.5 Soil information systems

Soil information systems (SISs) have become central components of modern soil science, driven by the growing demand for accessible, harmonised, and policy-relevant soil information (Arrouays et al., 2017b; Bouma, 2014; Omuto et al., 2024). A SIS is defined as an integrated platform that facilitates the storage, management, analysis, and dissemination of soil data and

associated environmental information (van Egmond et al., 2023). Unlike standalone databases or static soil maps, a functional SIS links multiple data types and provides mechanisms for querying, visualisation, and reuse by a wide range of users (Batjes et al., 2017; van Egmond et al., 2023).

Whereas historic soil surveys and maps were primarily produced for specific projects or sectors, land management, agricultural planning, climate reporting, and environmental monitoring increasingly require soil information that is standardised, georeferenced, spatially explicit, and capable of being queried, analysed, and updated as new data become available (McBratney et al., 2003; van Egmond et al., 2023). SISs are intended to meet these demands by integrating point observations (such as soil profile descriptions and laboratory analyses), spatial coverages (such as conventional and digital soil maps, terrain attributes, climate data, and remote-sensing products), metadata, and analytical tools within an internally consistent and logically organised digital structure (Omuto et al., 2012; van Egmond et al., 2023).

Soil information is increasingly recognised as a public good and SISs are expected to adhere to open-data principles that promote transparency, reproducibility, and long-term data stewardship exemplified by the FAIR principles — Findable, Accessible, Interoperable, and Reusable (van Egmond et al., 2023; Wilkinson et al., 2016).

The development and implementation of SISs are uneven across regions and are characterised by persistent conceptual and practical challenges, data unavailability and inaccessibility. In many countries, soil data exist but are uncatalogued and unpublished, spread across institutions, stored in incompatible formats, or restricted by legal, institutional, or personal barriers (Omuto et al., 2012; Rossiter, 2008). Incomplete metadata and undocumented analytical methods or provenance further limit data usability and comparability (Batjes, 2008; Batjes et al., 2017; Carvalho Ribeiro et al., 2020). As a result, the mere existence of soil data does not automatically translate into usable or decision-relevant soil information.

Standardisation and harmonisation constitute another major challenge for SIS development. International harmonisation is essential but must be balanced with local realities. Soil datasets are inherently heterogeneous, reflecting differences in survey objectives, sampling designs, classification systems, laboratory methods, and depth conventions (Batjes et al., 2020; Nachtergaele et al., 2012). SISs must therefore reconcile unstandardised legacy data with contemporary international standards and formats to enable modelling, comparison, and integration into continental and global initiatives (Arrouays et al., 2017b, 2020a). This reconciliation often demands trade-offs between preserving original observations and producing internally consistent datasets, making transparency, documentation, and the retention of raw data essential components of SIS design (Carvalho Ribeiro et al., 2020; Omuto et al., 2012).

Another challenge relates to sustainability and institutional context. SISs are not purely technical products. They depend on stable institutional mandates, skilled personnel, secure hosting environments, and long-term financial support (Omuto et al., 2012, 2024; van Egmond et al., 2023). Many SIS initiatives lose momentum after initial project phases due to unclear ownership, insufficient maintenance resources, or weak integration into governance and planning structures (Omuto et al., 2012, 2024). These challenges are particularly pronounced in countries where soil institutions are under-funded, under-staffed, and under-equipped, and where soil information competes with other national priorities. Consequently, global or continental SIS frameworks cannot simply be replicated at national level without careful consideration of local institutional realities and capacities (Omuto et al., 2024).

Recent conceptual work emphasises that SIS development should follow strategies linking user needs, data acquisition, data management, modelling, and dissemination (van Egmond et al., 2023). SISs are inherently iterative: new observations and modelling outputs feed back into system refinement, while user requirements evolve as information becomes available (Rossiter, 2008; van Egmond et al., 2023). Effective SIS implementation must therefore be rooted in a national context that reflects existing data legacies, institutional arrangements, technical capacity, and long-term governance structures (Omuto et al., 2012, 2024).

These considerations are particularly relevant in countries with sparse or fragmented soil data. In such contexts, SIS development cannot rely primarily on continuous inflows of new field data but must instead prioritise the rescue, harmonisation, and revitalisation of legacy soil information (Arrouays et al., 2017b; Rossiter, 2008). Conventional soil maps and expert knowledge often play a central role in structuring databases, guiding harmonisation decisions, and framing digital modelling outputs (Arrouays et al., 2017b, 2020a; Walter et al., 2006). In data-scarce environments, SISs must therefore integrate conventional and digital approaches rather than treating them as sequential or substitutable.

3.6 Legacy soil data

Legacy soil data refer to the accumulated body of soil information generated through past surveys, mapping campaigns, and research projects, including soil maps, profile descriptions, laboratory analyses, field notes, photographs, land evaluation reports, and associated metadata (Batjes, 2008; Minai et al., 2022; Rossiter, 2008). These include observed characteristics of sites, profiles, land surfaces, and soil forming factors. In addition to soil information, legacy data may cover information on the local climate, geology, geomorphology, land quality, crop suitability, land use, and farming systems. The legacy data can be in the form of printed or digital documents, maps and images, tables, spreadsheets, or databases (Batjes, 2008; Omuto et al., 2024; Rossiter, 2008). In many regions, these datasets represent the primary field observation and laboratory analysis records of soil conditions and processes (Omuto et al., 2012). Despite their central importance, legacy soil data are frequently underutilised, poorly documented, and at risk of permanent loss (Arrouays et al., 2017b; Rossiter, 2008).

The loss of legacy soil information represents an irreversible erosion of scientific knowledge. Soil observations cannot be retrospectively recreated once sampling sites are lost, landscapes are transformed, or institutional memory fades (MacMillan et al., 2005; Rossiter, 2008). Some of the justifications for preserving legacy soil information are the labour and resources that were expended in their compilation, their value as inputs for digital mapping, and as baselines for temporal change detection and monitoring (Arrouays et al., 2017b; Baxter & Crawford, 2008). Their preservation is not merely a technical task but a form of scientific risk management, aimed at safeguarding long-term knowledge of soil resources (Omuto et al., 2012). Recognising this, legacy data rescue is seen as a fundamental activity in SIS development and digital soil mapping (Arrouays et al., 2017b, 2020c; Minai et al., 2022; Rossiter, 2008).

In Namibia, no comprehensive, state-funded soil survey has been conducted recently or is planned for the foreseeable future. Consequently, historical soil data form the backbone of national soil knowledge. These data have been fragmented across institutions, stored in analogue formats, or documented only in grey literature, limiting accessibility and long-term preservation. Regional and international initiatives demonstrate how rescued legacy data can be integrated into quality-

assessed, standardised platforms that support both research and policy applications when conducted systematically and transparently (Batjes et al., 2017, 2020; Leenaars et al., 2014). The consolidation of Namibia's dispersed legacy soil information is, therefore, a prerequisite for developing a functional national soil information system, producing an updated national soil map, and enabling reliable digital soil mapping.

A defining characteristic of legacy soil data is heterogeneity, which poses challenges for data integration (Carvalho Ribeiro et al., 2020; Dobos et al., 2010). Historical soil surveys were conducted for diverse purposes — such as irrigation planning, agricultural development, or engineering feasibility — using evolving classification systems, and different sampling designs, laboratory methods, and depth conventions (Arrouays, et al., 2017b). Spatial referencing may be incomplete or imprecise, analytical protocols may no longer be standard, and metadata are often sparse or absent (Batjes, 2008; Batjes et al., 2024; Carvalho Ribeiro et al., 2020). These attributes complicate direct reuse and pose challenges for integration with modern spatial datasets and models (McBratney et al., 2003). Nevertheless, in data-scarce environments, legacy datasets are precious: the cost and logistical complexity of repeating national-scale soil surveys make it impractical to replace historical observations with entirely new data (Arrouays et al., 2017b). Preserving and repurposing existing soil information is both economically and practically justifiable.

Legacy soil data revitalisation, also referred to as data *rejuvenation* (Minai et al., 2022; Rossiter, 2008) or data *resurrection* (Dent & Ahmed, 1995), comprises three interrelated processes:

3.6.1 Data archaeology

The first, often termed *data archaeology*, involves locating, cataloguing, and documenting dispersed soil information and reconstructing its provenance (Arrouays et al., 2017b; Odeh et al., 2012; Rossiter, 2008). This process involves searches in institutional archives, personal collections, and grey literature, and requires interpretation of undocumented conventions used by earlier surveyors (Omuto et al., 2012). Data archaeology establishes the context for subsequent quality assessment and harmonisation.

3.6.2 Data rescue

The second process, *data rescue*, focuses on stabilising and digitising vulnerable materials. This includes scanning analogue maps and documents, transcribing field forms and laboratory sheets, georeferencing spatial information, converting disparate digital formats into structured databases (Odeh et al., 2012; Omuto et al., 2012), and storing them in stable digital repositories. Data rescue also involves preliminary quality control such as identifying duplicate records, resolving obvious inconsistencies, and documenting uncertainty where original information is incomplete (Batjes et al., 2017; Odeh et al., 2012; Omuto et al., 2012). Effective data rescue preserves raw observations alongside derived or harmonised values, allowing future users to reassess assumptions and methodologies (Omuto et al., 2012). Original, non-interpreted field observations allow for the derivation of new soil property layers and more flexible integration across datasets collected under different systems (Dobos et al., 2010).

3.6.3 Data renewal

The third data revitalisation process, *data renewal*, addresses the integration of rescued data into contemporary classification, analytical, geospatial and modelling frameworks, to ensure cross-platform compatibility with international repositories. Renewal typically requires harmonisation of soil classifications to current systems such as the World Reference Base for Soil Resources, WRB (IUSS Working Group WRB, 2015, 2022)), standardisation of analytical units and depth intervals, and alignment with widely used coordinate reference systems (Batjes et al., 2020; Cambule, 2013; Cambule et al., 2015; Leenaars et al., 2014). Depth harmonisation, uncertainty estimation, and linkage with environmental covariates are central components of this process, allowing legacy observations to contribute meaningfully to digital soil mapping and soil function assessment (Arrouays et al., 2020c; McBratney et al., 2003). Several authors emphasise the need for geodetic control of both point and areal GIS coverages, the use of adequacy criteria to assess legacy data quality, incorporation of satellite imagery, environmental, and terrain parameters, comprehensive metadata compilation, and integration of renewed data into accessible spatial data infrastructures (Cambule, 2013; Cambule et al., 2015; Rossiter, 2008). Data renewal transforms static historical records into dynamic resources capable of supporting current and future soil information needs.

While legacy data revitalisation offers great benefits, it also introduces methodological trade-offs. Harmonisation inevitably involves simplification and the potential loss of locally specific detail, while translation between classification systems may introduce ambiguity (Arrouays et al., 2017b). It is, thus, essential to provide comprehensive metadata and preserve original data alongside harmonised outputs (Omuto et al., 2012). Rather than disguising uncertainty, effective revitalisation strategies document it honestly, allowing users to judge fitness for purpose (Rossiter, 2018).

3.7 Conventional soil mapping (CSM)

Conventional soil maps are core components of national soil information systems. Expert-derived, polygon-based soil maps embody integrated understanding of soil–landscape relationships, parent material controls, and pedogenic history.

3.7.1 Conceptual CSM framework

In CSM, soil surveyors construct qualitative mental models to understand the co-evolution of landscapes and soils (Bui, 2004; Hudson, 1992; Soil Survey Staff, 2017). These abstract frameworks integrate multiple scales of observation (from soil profile point data to satellite images) and pedogenic reasoning (an understanding of chemical, physical and biological processes in soil genesis, and extrapolations from comparable sites) to predict soil distribution patterns and create soil maps (Bui, 2004; Hudson, 1992; Ma et al., 2019; MacMillan et al., 2005). CSM is thus a conceptual interpretation based on field-observed and laboratory-measured data, expert insight, years of accumulated experience, intuitive reasoning, and the ability to synthesise complex landscape information into clear, well-defined spatial representations that cannot be fully inferred from statistical models alone (Hudson, 1992; Ma et al., 2019; MacMillan et al., 2005; Soil Survey Staff, 2017). These maps reflect both scientific training and professional artistry in pattern

recognition (McBratney et al., 2025).

3.7.2 Methodological considerations in synthesising a national map from legacy surveys

The synthesis of a national soil map from existing maps presents methodological challenges due to differences in survey purpose, sampling intensity, mapping scale, methodology, and classification systems (Batjes et al., 2024; Nachtergaele et al., 2012). The integration of such heterogeneous soil maps requires harmonisation of both spatial and attribute information.

Conventional soil maps are based on a soil–landscape paradigm and represent the soil mantle as discrete units (pedons or polypedons) that are interpretable at a given cartographic scale, rather than as continuous gradients (Bui, 2004; Burrough, 1989; Hudson, 1992). Soil map units function as conceptual entities that group soils expected to co-occur within similar environmental settings, based on pedological understanding, field observations, and extrapolation from comparable sites. As such, map units¹ represent dominant, associated and included soils, each identified by a distinct typological label², within a single polygon (Bui, 2004; Burrough, 1989; Hiederer et al., 2013; Hudson, 1992).

Map unit concepts — such as consociations, associations, complexes, and phases — may differ among surveys and therefore require standardisation to ensure comparable interpretation of soil composition and within-unit variability (Soil Survey Staff, 2017).

Legacy surveys often employ national classification schemes, earlier versions of international systems, or project-specific taxonomies. These have to be translated into a common reference classification, such as the WRB, using correlation tables and expert judgement (IUSS Working Group WRB, 2015; Zádorová et al., 2021).

To ensure spatial consistency across the national compilation, differences in mapping scale require that small-scale map units be subdivided, while those of large-scale maps need to be generalised or aggregated to match the target resolution (Nachtergaele et al., 2012). Edge mismatches and discontinuities at survey boundaries must be resolved using ancillary data such as terrain, attributes, remote sensing imagery, and geological information. The minimum mappable unit is determined by the intended final map scale, landscape complexity, and density of supporting observations, thus smaller soil bodies may be represented implicitly through map unit composition rather than explicit delineation (FAO, 2006; Soil Survey Staff, 2017).

Uncertainties arising from variable reliability of source maps, differences in survey methods, and assumptions introduced during harmonisation must be documented to transparently support interpretation of compiled soil maps, spatial analysis, and subsequent modelling (Burrough, 1989; MacMillan et al., 2005).

¹ Soil Map Unit (SMU): A spatial unit of a soil map that can consist of a single STU, but more commonly is a combination of several STUs (dominant, associated and/or included soils) occurring within the mapped area. Each SMU has a unique label or code (which can be a combination of the labels/codes of its constituent STUs). A SMU can consist of several non-contiguous polygons in a vector soil map (adapted from Hiederer, 2013).

² Soil Typological Unit (STU): A distinct, homogeneous soil type that can be described with a specific set of characteristics (e.g. soil horizon sequence, texture, structure, pH, etc.) and classified in the WRB in terms of a Reference Soil Group and qualifiers. It can be either a dominant, associated or included soil type within a soil map unit (adapted from Hiederer, 2013).

3.7.3 Strengths and limitations of CSM

Categorical soil maps are well-established and widely understood tools for communicating soil information. They are extensively used in land evaluation, agriculture, and environmental assessment because they integrate multiple soil properties into classes and map units that reflect functional soil behaviour (e.g. drainage, rooting depth, fertility constraints) (FAO, 2006; Soil Survey Staff, 2017). Map units are typically delineated according to natural landscape organisation, with boundary placement being informed by geomorphic and environmental transitions, which assists interpretation (Hudson, 1992; van Engelen & Dijkshoorn, 2013). Expert knowledge confers a high degree of pedological realism and interpretability, such that the resulting maps are meaningful rather than purely statistically derived (McBratney et al., 2003; van Engelen & Dijkshoorn, 2013).

CSM has some recognised limitations. Field surveys and laboratory analyses are time-consuming, labour-intensive, and costly (Grunwald et al., 2011; Kempen et al., 2012). In addition, practical constraints such as difficult terrain, access restrictions, long travel distances, and socio-political factors can limit field data collection and contribute to uneven spatial coverage (FAO, 2006; Soil Survey Staff, 2017).

A degree of subjectivity is inherent in the mapping process. This may arise in boundary delineation (positional uncertainty), map-unit composition (within-unit variability and purity), soil classification (thematic uncertainty), and the interpretation of auxiliary information such as aerial imagery or terrain models (Burrough, 1989, 1991; Hudson, 1992; Minasny & McBratney, 2016). As a result, soil maps are not exactly reproducible by different surveyors working independently, particularly in transitional landscapes or where soil patterns are subtle (Bazaglia Filho et al., 2013; Han et al., 2022). While such subjectivity can be mitigated through correlation rules, standardised legends, field verification, peer review, and detailed documentation, it cannot be eliminated entirely (Bui et al., 2020; Hudson, 1992).

Conventional soil maps are inherently categorical, representing soils as discrete polygons. This can imply internal homogeneity and sharp boundaries that do not reflect the continuous nature of soil variation (Burrough, 1989, 1991; Hudson, 1992; Minasny & McBratney, 2016). Map legends and attribute tables can only capture a subset of all the information known to the surveyor, meaning that soil maps are necessarily simplifications of reality and cannot wholly encode tacit expert knowledge (Bui et al., 2020; Dobos et al., 2010; MacMillan et al., 2005). Conventional soil maps also represent soils at a specific point in time and generally do not capture temporal dynamics and land-use change effects (McBratney et al., 2003). Interpretation often requires some familiarity with soil classification systems, which may limit accessibility for non-specialist users (Bui et al., 2020; Hudson, 1992).

CSM is constrained by scale. National- and regional-scale maps are not suitable for field-level decision-making or precision agriculture applications. Soil bodies smaller than the minimum mappable area are typically omitted or included as minor components within map units, with their spatial location remaining uncertain (FAO, 2006; Soil Survey Staff, 2017).

These characteristics hamper the direct integration of conventional polygon-based outputs into modern grid-based geospatial tools and models (Arrouays et al., 2020a).

Despite these limitations, CSM remains indispensable, particularly at national and regional scales

and in data-scarce environments. By synthesising expert knowledge, soil–landscape relationships, and field observations into coherent spatial units, conventional maps provide an interpretive framework for land evaluation and digital soil mapping outputs (Bui, 2004; Bui et al., 2020; Grunwald et al., 2011). CSM can support DSM in several areas, such as showing the relative distribution of soil, identifying soil forming factors, demonstrating correlation and spatial structures of soil properties, and providing information on the temporal dynamics of biochemical and physical processes (Walter et al., 2006). Rather than being superseded by DSM, conventional mapping complements it: providing context for training data and reference information for validation; as a covariate input; and constraining model predictions (Rasaei et al., 2020). CSM thus forms a structural backbone within which gridded predictions can be interpreted in pedological terms (Arrouays et al., 2020a; Rossiter & Poggio, 2025).

3.8 Digital soil mapping (DSM)

Some limitations of CSM were addressed by the development and adoption of predictive or digital soil mapping (DSM), a field that has gained momentum since it was formalised and quantified in 2003 (McBratney et al., 2003; Minasny & McBratney, 2016).

DSM developed from pedometrics — the application of mathematical and statistical methods to study the distribution, characteristics, and genesis of soils (Heuvelink, 2003; Wadoux et al., 2023) — and focuses on developing methods for automated prediction of soil classes or properties at user-defined resolutions (Lagacherie & McBratney, 2007). DSM has been defined as “*the generation of geographically referenced soil databases based on quantitative relationships between spatially explicit environmental data and measurements made in the field and laboratory*” (McBratney et al., 2003), and as “*the creation and population of spatial soil information systems by numerical models inferring the spatial and temporal variations of soil types and soil properties from soil observation and knowledge and from related environmental variables*” (Lagacherie & McBratney, 2007).

3.8.1 Conceptual DSM frameworks

DSM is rooted in foundational soil science theories, including Dokuchaev's (1883) early work on soil formation, Jenny's original and Integrated CLORPT models (Jenny, 1941; Jenny et al., 1968), the SCORPAN model (McBratney et al., 2003), and the STEP–AWBH model (Grunwald et al., 2011; Thompson et al., 2012). These approaches built upon topofunctions (Walker et al., 1968), pedotransfer functions (Wösten, 1997), pedometrics (Burrough et al., 1994; Webster, 1994) and digital soil morphometrics (Hartemink & Minasny, 2014), and advances in geographic information systems (Burrough, 1986, 1991), digital terrain analysis (Gallant & Wilson, 1996), geostatistical soil modelling (Webster, 2000), and machine learning (Bui & Moran, 2003). Further impetus was provided by global initiatives like GlobalSoilMap (Arrouays et al., 2014, 2017b, 2020b, 2020c) and SoilGrids (Hengl et al., 2014, 2015, 2017a, 2017b).

In the *SCORPAN model* (Figure 9.), soil classes or properties at unobserved locations are inferred using a deterministic function involving previously measured soil attributes (S), climatic properties (C), biotic activity (O), terrain features (R), parent material (P), soil age or stage of development (A), and neighbourhood context (N), with a spatially dependent residual term (ϵ) that accounts for unexplained variation (McBratney et al., 2003; Yigini et al., 2018). The model

quantifies the relationships between measured soil (point) data and environmental covariates, allowing for the spatially continuous prediction of soil properties and classes across the mapping area, and estimation of prediction uncertainty (Kienast-Brown et al., 2017; Zhang et al., 2017).

$Sa_{[x,y,t]}$	$= f(S_{[x,y,t]}, C_{[x,y,t]}, O_{[x,y,t]}, R_{[x,y,t]}, P_{[x,y,t]}, A_{[x,y,t]})$
Sa	= soil attribute of interest
S	= other soil properties / from nearby
C	= climate / climatic properties
O	= organisms, including fauna, flora, human activity
R	= relief / topography / landscape attributes
P	= parent material / lithology
A	= substrate age / degree of soil development
N	= spatial position / coordinates of a specific location
x; y	= coordinates of field observations
t	= time frame when soil data were collected (ideally the same as when the soil forming factors were active)

Figure 9. The SCORPAN model (McBratney et al., 2003; Yigini et al., 2018).

The *STEP–AWBH model* (Grunwald et al., 2011; Thompson et al., 2012) extends the SCORPAN model to reflect the dynamic nature of soil systems. This model integrates ancillary soil properties (S), topographic properties (T) (e.g. elevation, slope gradient, slope curvature and compound topographic index), ecological properties (E) (e.g. physiographic regions or ecoregions), parent material or geologic properties (P) (e.g. geologic formations), atmospheric properties (A) (e.g. precipitation, temperature and solar radiation), water properties (W) (e.g. soil moisture and surface runoff), biotic properties (B) (e.g. vegetation, land cover, spectral indices and organisms), and human-induced forcings (H) (e.g. land use, land use change, contamination and disturbances). The static environmental factors (STEP) are thus augmented by dynamic factors (AWBH) that enable simulations of soil transformation over time (Zhang et al., 2017).

These two models guided the selection of covariates and the interpretation of model outputs in this study.

3.8.2 Machine-learning approaches in predictive soil mapping

In essence, DSM uses geostatistical or machine-learning (ML) techniques to model complex, nonlinear relationships between known soil data and environmental variables to make spatially continuous predictions of soil properties or soil classes over an area of interest, using large spatial datasets (remote sensing, terrain models, climate surfaces), field observations, and laboratory analyses (Lagacherie & McBratney, 2007; McBratney et al., 2003).

ML is a subset of artificial intelligence that involves the development of algorithms and statistical models that enable computers to improve their performance on a specific task by learning patterns, relationships and rules from examples (the training dataset) and experience without being explicitly programmed, and then making predictions or decisions based on the learnt information (Bishop, 2006; Mitchell, 1997). During model training, internal parameters of the algorithms are optimised to minimise the difference between model predictions and actual target values, whereafter model performance is evaluated (validated) on unseen data (the testing dataset) to assess its generalisation capability (Bishop, 2006; Mitchell, 1997).

ML algorithms can operate effectively with complex datasets and large numbers of predictors, even when relationships are poorly specified and covariates are cross correlated (Bishop, 2006; Mitchell, 1997). They do not require transformation of the observational data because they do not

make assumptions about the data distribution (Wadoux et al., 2020). ML techniques are generally grouped into linear-based, kernel-based, probabilistic-based, and tree-based approaches (Yigini et al., 2018).

DSM shifts the representation of soil from discrete polygons to continuous surfaces — raster maps of two-dimensional grid cells — with predicted numeric attribute values (e.g. % soil organic carbon) or classes (e.g. soil types). It allows integration of multiple soil properties across depths, quantitative assessment of model accuracy, quantification of prediction uncertainty, reproducible modelling methods, and ease in updating models as new data become available. DSM is fast, cost-effective and embeds information on the relative contributions of various pedogenic factors on soil attribute distribution (Arrouays et al., 2020a).

DSM has become the global standard for modern soil information systems and constitutes the core predictive component of NamSIS.

3.8.3 Methodological considerations in DSM

Beyond the conceptual framework, practical implementation of DSM requires a number of methodological decisions. Seven aspects are particularly important for this study: (1) Quality of soil data; (2) choice of machine learning algorithm; (3) modelling environment; (4) covariates; (5) model tuning, training and validation; (6) performance metrics; and (7) uncertainty estimation.

3.8.3.1 Quality of soil data

The quality and representativeness of soil observations are important determinants of DSM performance. Soil data used for model calibration should ideally provide adequate spatial coverage across the environmental gradients of the study area, while avoiding strong clustering or systematic sampling bias that could distort model predictions (Hengl et al., 2003; McBratney et al., 2003). Positional accuracy, consistency of laboratory methods, and completeness of metadata are also important, as errors and inconsistencies may propagate through the modelling process (Nenkam et al., 2024). Legacy datasets often contain heterogeneous analytical methods, variable sampling depths, and differing classification systems, requiring harmonisation before integration into DSM workflows (Arrouays et al., 2014; Malone et al., 2017). Soil observations should also be matched as closely as possible to the spatial resolution and temporal context of the environmental covariates used for modelling, since scale mismatches can introduce additional uncertainty (Minasny & McBratney, 2016; Wadoux et al., 2020). In arid and semi-arid environments, where soil properties may exhibit narrow ranges and strong local variability, careful consideration of data quality, representativeness, and uncertainty is particularly important.

3.8.3.2 Choice of machine learning algorithm

Early DSM applications relied primarily on linear regression, regression kriging, and geostatistical approaches, but these methods often struggled to capture the nonlinear and interacting effects of climate, terrain, vegetation, and parent material on soil properties. As a result, a wide range of machine-learning (ML) algorithms have been adopted in DSM, including artificial neural networks, Support Vector Machines, Boosted Regression Trees, Gradient Boosting Machines, and Random Forests (Behrens et al., 2018; McKenzie & Ryan, 1999; Wadoux et al., 2020). The

selection of an appropriate approach requires trade-offs: geostatistics handles spatial autocorrelation explicitly and provides interpretable variograms but often relies on assumptions of linearity and normality and may struggle with high-dimensional covariate sets (Wadoux et al., 2020). Machine learning methods, on the other hand, handle non-linear relationships and complex interactions more flexibly but can be harder to interpret and require careful tuning and validation (Wadoux et al., 2020). Benchmarking studies have shown that the performance of ML algorithms is dependent on context and the dataset, thus there is no single universally best algorithm (Schmidinger et al., 2025). Tree-based ensemble methods are popular due to their flexibility, robustness, strong predictive performance across soil properties and environments, and interpretability (Wadoux et al., 2020).

Random Forests (RF) is a non-parametric ensemble method that builds multiple decision trees using bootstrap samples of the training data (Breiman, 2001; Efron & Tibshirani, 1994). Predictions are obtained by averaging the outputs of individual trees (for regression) or by majority vote (for classification). RF can handle non-linear relationships, mixed data types, and high-dimensional predictor sets without strong statistical assumptions, and is generally tolerant of noise, outliers and missing values (Heung et al., 2014). RF provides internal estimates of variable importance, and it benefits from hyperparameter tuning to enhance performance (Liaw & Wiener, 2002). Its suitability for DSM is well established, with numerous studies reporting strong predictive accuracy across a range of soil properties (Forkuor et al., 2017; Grimm et al., 2008; Hengl et al., 2015; Kumar et al., 2023; Vandana et al., 2024; Wiesmeier et al., 2011). Comparative reviews show that RF often outperforms other machine learning algorithms and traditional regression models (Lamichhane et al., 2019; Mgohele et al., 2024). Using soil observations from Namibia, Van der Westhuizen et al. (2022) evaluated RF and other modelling approaches for digital prediction of topsoil clay content, with particular emphasis on accounting for uncertainty in the calibration data. They found that their proposed measurement-error-filtered methods did not improve on RF prediction accuracy in Namibia, largely because the estimated measurement errors were relatively small and did not vary enough. The only other DSM study known from Namibia, Nijbroek et al. (2018), used field soil observations from Otjozondjupa Region to establish a soil organic carbon baseline for monitoring Land Degradation Neutrality. They concluded that RF kriging was statistically the best-performing approach when relatively few local samples were available, and that more sophisticated ML does not automatically translate into better maps. Although more sophisticated ensemble and hybrid approaches may offer further improvements in some contexts (Guo et al., 2015; Hengl et al., 2017b, 2021), RF remains a practical, computationally efficient, and widely adopted option for national- and regional-scale DSM (Brungard et al., 2015; Heung et al., 2014, 2016; Wadoux et al., 2020).

3.8.3.3 Choice of modelling environment

Google Earth Engine (GEE) is a cloud-based geospatial environment designed for large-scale analysis of satellite and environmental data (Gorelick et al., 2017), making it highly suitable to meet the heavy computational demands of national-level DSM. GEE provides direct access to extensive remote sensing, climatic and terrain datasets through the Earth Engine Data Catalogue. It includes a built-in RF algorithm, supports on-platform computation of derivatives such as vegetation indices and terrain attributes, and offers spatial operations like masking, resampling, and raster overlay. Previous DSM studies have demonstrated the benefits of GEE's ready-to-use covariates and rapid parallel processing capabilities through Google's servers (Padarian et al.,

2015). However, GEE allows relatively limited control over model tuning and customisation. Conversely, *R* (R Core Team, 2024a, 2024b) supports a wide range of ML algorithms, customised operations, and better control over model parameters and visualisation, but is constrained by local processing capacity during large-scale spatial modelling.

3.8.3.4 Covariate selection

DSM applications often start with a large set of potential environmental covariates representing the soil forming factors, derived from digital elevation models, satellite imagery, climate surfaces, and geological and land cover datasets. The number varies substantially with mapping extent, available datasets and the feature-selection strategy. Relatively focused studies have used as few as five to seven environmental covariates (Padarian et al., 2015; Samuel-Rosa et al., 2015), whereas continental and global studies have started with much larger predictor stacks. SoilGrids1km used 75 original environmental covariates (Hengl et al., 2014), AfSoilGrids250m used approximately 45–111 predictors depending on the target property (Hengl et al., 2015), and SoilGrids250m used 158 covariates (Hengl et al., 2017b). Nussbaum et al. (2018) evaluated candidate sets of approximately 300–500 covariates, depending on the study region, but retained only about 3–6% in models. Similarly, SoilGrids 2.0 reduced more than 400 initial environmental layers to approximately 150 after correlation filtering and ultimately retained 20–40 predictors per soil-property model (Poggio et al., 2021). Kasraei et al. (2024) reduced 70 initial covariates to property-specific subsets of only 4–12 variables with little or no loss of predictive performance. These studies show that large initial covariate libraries are common in contemporary digital soil mapping, but that the number retained in the final models is generally property-specific and can range from fewer than 10 to several tens of predictors.

Various studies have emphasised the value of reducing the predictor set to improve computational efficiency, avoid multicollinearity, reduce the risk of overfitting, and obtain more reliable estimates of variable importance (Gomes et al., 2019; Wadoux et al., 2020). Streamlining a covariate set reduces model complexity and facilitates clearer interpretation of soil–environment relationships and pedogenic processes (Kasraei et al., 2024; Ma et al., 2019; Nussbaum et al., 2018). A parsimonious predictor set does not necessarily compromise predictive accuracy (Kasraei et al., 2024; Samuel-Rosa et al., 2015). Covariate selection can be done through correlation-based screening or by using wrapper methods (e.g. recursive feature elimination) that iteratively remove the least important predictors based on model performance (Gomes et al., 2019; Kasraei et al., 2024; Kuhn & Johnson, 2013; Wadoux et al., 2020).

Wadoux et al. (2020) stressed the importance of including pedological knowledge — understanding soil as they occur in their environment — in the selection of DSM predictors and being able to interpret the results in terms of the underlying soil processes.

3.8.3.5 Model tuning, training and validation

Machine-learning algorithms such as Random Forests, Gradient Boosting Machines, and Support Vector Machines require the specification of hyperparameters that control model complexity and learning behaviour, including tree depth, number of trees, minimum node size, and sampling strategies (Kuhn & Johnson, 2013). Hyperparameter tuning is essential to balance bias and variance, and to prevent overfitting, particularly when working with high-dimensional covariate

sets and unevenly distributed legacy soil data (Probst et al., 2019a, 2019b; Wadoux et al., 2020). Common tuning approaches include manual search, grid search, which exhaustively evaluates predefined parameter combinations, and random search, which samples the parameter space probabilistically and is often more efficient in high-dimensional settings (Bergstra & Bengio, 2012). Model-based (a.k.a. Bayesian) optimisation approaches are increasingly adopted in large DSM operations, as they iteratively learn the relationship between hyperparameters and model performance using surrogate models and propose increasingly promising parameter combinations (Probst et al., 2019a, 2019b; Wadoux et al., 2020).

To assess model validity, datasets are usually partitioned into calibration (training) and validation (testing) sets, using fixed random or stratified splits, k-fold cross-validation, or bootstrap resampling (Chen et al., 2022). Choice of strategy depends on sample size, spatial clustering, computational constraints, and the intended application of the model. An 80:20 (or 70:30) split ensures that a sufficiently large proportion of the data is used for model calibration to represent important soil–landscape relations, while reserving an adequate independent subset for performance evaluation (Liu et al., 2022; Samuel-Rosa et al., 2015). Applying multiple bootstrap iterations with replacement ensures more defensible estimates of model accuracy and better accounts for sampling variability in legacy soil datasets with clustered data than a single split (Efron & Tibshirani, 1994; Padarian et al., 2015).

3.8.3.6 Performance metrics

In DSM studies, no single performance metric fully characterises model behaviour; therefore, multiple complementary statistics are typically reported to capture different aspects of model performance, including accuracy, precision, and agreement (Wadoux et al., 2020; Wadoux & Minasny, 2024).

The *coefficient of determination* (R^2) quantifies the “fitting potential”, the proportion of variance in the observed data that is explained by the model. It is a measure of precision, but it does not indicate whether predictions are unbiased or accurate (Kuhn & Johnson, 2013). As it is sensitive to the range of observed values, R^2 is often low in low-variance datasets — typical of arid-zone soil data — even when model performance is otherwise acceptable (Hastie et al., 2009; Legates & McCabe, 1999).

The *root mean squared error* (RMSE) reflects overall predictive accuracy in terms of bias (systematic over- or underprediction) and variance (sensitivity to sampling variability and noise) in model predictions and quantifying the magnitude of residual errors (Chai & Draxler, 2014; Hastie et al., 2009). It represents the standard deviation of prediction errors and therefore describes the distribution of residuals around the fitted model. RMSE is intuitive because it shares the units of the target variable (Chai & Draxler, 2014). Lower RMSE values indicate better performance but must be interpreted relative to the natural variability and measurement uncertainty of the soil property (Legates & McCabe, 1999; Moriasi et al., 2007). RMSE is particularly appropriate when absolute prediction error matters more than explained variance, when property ranges are narrow — as is often the case in arid soils — and when models are compared across products (Chai & Draxler, 2014). Model stability can be assessed by bootstrap resampling, where a narrow distribution of RMSE values indicates good and consistent predictive performance (Efron & Tibshirani, 1994; Hastie et al., 2009).

Lin’s *concordance correlation coefficient* (CCC) assesses the agreement between two continuous

variables (observed vs. predicted values, in this case) by jointly accounting for accuracy (deviation from the 1:1 line in a scatterplot) and precision (correlation), with 1 equalling perfect agreement, 0 meaning no agreement beyond chance, and <0 indicating systematic disagreement (Lin, 1989).

In DSM, what constitutes good performance is context-dependent and varies by soil property, spatial scale, and data availability (see Appendix T – Summary of model performance metrics reported in selected digital soil mapping studies). National-scale studies typically report R^2 values of ~0.3–0.6 for many soil properties, with RMSE values comparable to or lower than analytical uncertainty, and CCC values above ~0.4–0.6 considered acceptable (Hengl et al., 2017a, 2017b; Poggio et al., 2021; Wadoux et al., 2020). Model performance should not be judged on metrics alone; it must also be interpreted in terms of uncertainty estimates and pedological plausibility (Grunwald, 2009; Malone et al., 2017; Poggio et al., 2021).

3.8.3.7 Uncertainty estimation

Contemporary DSM emphasises the quantification and communication of spatial prediction uncertainty, often in the form of prediction intervals. Uncertainty can be estimated through several approaches, including *bootstrapping* of training data, analysis of residual variance, Quantile Regression Forests, or ensemble modelling (Heuvelink, 2018; Wadoux et al., 2020). Whereas some uncertainty methods are model-specific (e.g. kriging, Quantile Regression Forest), bootstrap-based ensemble methods are attractive because they do not assume normal error distributions, require access to internal model coefficients, gradients, etc. or rely on analytical expressions of prediction variance (Breiman, 2001; Efron & Tibshirani, 1994). Bootstrap methods capture both sampling variability and model instability and are widely used with linear regression, Random Forests, Gradient Boosting, Support Vector Machines and neural networks (Efron & Tibshirani, 1994; Padarian et al., 2015).

The *Prediction Interval Coverage Probability (PICP)* is commonly used as a diagnostic of uncertainty calibration, by measuring the proportion of observed values falling within a specified prediction interval. Well-calibrated models should produce PICP values close to the nominal interval probability (e.g. ~90% for a 90% interval), while systematic deviations indicate under- or overconfident models (Poggio et al., 2021).

3.8.4 Strengths, limitations, knowledge gaps, and current debates

DSM is a powerful approach for producing quantitative, high-resolution, and updateable soil datasets. A key strength is that it generates quantitative estimates of soil properties that can be directly used in a wide range of modern geospatial analyses, models, and decision-support systems (Grunwald, 2009; Malone et al., 2017). DSM products are spatially continuous, typically represented as raster grids, thereby avoiding the assumption of internal homogeneity implied by discrete map units of conventional soil maps (McBratney et al., 2003). The increasing availability of high-resolution remote sensing, terrain, and climate datasets enables DSM to cover large areas consistently and, in many cases, at finer spatial resolution than conventional maps, including scales relevant to field-level applications. Kempen et al. (2012) demonstrated that DSM can be more cost-effective than CSM over large areas, provided that sufficient legacy soil data and high-quality environmental covariates are available.

DSM approaches are generally repeatable and updateable, as they are based on explicit input data,

workflows, and algorithms that can be re-run to confirm previous results or to incorporate new data and improve predictions over time (Arrouays et al., 2014; McBratney et al., 2003; Minasny & McBratney, 2016). Many DSM applications provide measures of prediction uncertainty, which are important for assessing the reliability of outputs and guiding future data collection (Malone et al., 2017; Wadoux et al., 2020)).

DSM is, however, also subject to several limitations. Model performance depends strongly on the availability, representativeness, and quality of soil observations used for calibration, and is negatively influenced by the variety of laboratory methods, missing data, and uneven profile distribution (Láng et al., 2016). In data-scarce regions, sparse and unevenly distributed samples can result in biased models and reduced predictive accuracy (Hengl et al., 2017a, 2021; McBratney et al., 2003). Nenkam et al. (2024) identified low sampling density, spatial clustering, and positional errors as major constraints in African soil datasets. They noted the continued reliance on legacy data with uneven coverage and the lack of probability-based validation sampling. In addition, most soil data represent the topsoil of agricultural land (McBratney et al., 2025). These limitations underscore the need for both legacy data rescue and targeted new data collection.

DSM relies on environmental covariates as proxies for soil-forming factors, and model performance depends on how effectively these covariates capture underlying soil–environment relationships. Improving understanding of which environmental variables best explain soil variability can support more process-informed covariate selection (Adeniyi et al., 2024). Predictions are also sensitive to the spatial resolution and quality of covariates, and mismatches between the scale of soil observations and environmental datasets can introduce additional uncertainty (Ma et al., 2019; Minasny & McBratney, 2016; Wadoux et al., 2020). In addition, many soil functions of interest — such as carbon sequestration, nutrient cycling, and water storage — are dynamic and are influenced by subsoil processes, highlighting the need for repeated observations, time-series covariates, and improved depth modelling approaches (Grunwald et al., 2011; Malone et al., 2009; Minasny & McBratney, 2016).

Global soil products such as SoilGrids have significantly improved access to soil information at continental and global scales but often lack the spatial detail and contextual accuracy needed for national applications (Buenemann et al., 2023; Nenkam et al., 2024). This reflects a broader distinction between global and locally calibrated DSM approaches. Global models benefit from large datasets and methodological consistency but tend to generalise across diverse environments, which can reduce local accuracy (Buenemann et al., 2023; Nijbroek et al., 2018; Stoorvogel et al., 2009). In contrast, national or regional DSM efforts that incorporate locally relevant covariates and legacy soil observations have been shown to improve predictive performance and better represent landscape-specific soil variability (Nenkam et al., 2024). This has led to increasing emphasis on context-specific, bottom-up modelling approaches, particularly in data-scarce regions (Arrouays et al., 2017a; Bohn & Miller, 2024; Mulder et al., 2016).

The use of advanced machine-learning algorithms, including ensemble methods and deep learning, has improved the ability of DSM to capture complex, non-linear relationships and integrate diverse datasets. However, these approaches can reduce interpretability. Models may function as “black boxes”, making it difficult to explain predictions in pedological terms, which can reduce user confidence and hinder practical application (Bui et al., 2020; Wadoux et al., 2020). DSM predictions do not inherently incorporate the pedological understanding embedded in conventional soil maps and may therefore not always align with known soil–landscape relationships, especially in complex terrain or where training data are limited. This underscores the continued need for

expert evaluation and the integration of conventional soil knowledge when interpreting DSM outputs (Arrouays et al., 2014; Bui, 2004; Bui et al., 2020).

While many DSM models produce uncertainty estimations, these typically reflect model uncertainty, not uncertainties arising from input data, measurement error, or preprocessing steps such as depth harmonisation (Kuhn & Johnson, 2013). Communicating the meaning and implications of uncertainty to end-users remains a recognised challenge (Heuvelink, 2018; Wadoux et al., 2020; Wadoux & Minasny, 2024).

DSM also requires a diverse skill set and often needs multidisciplinary collaboration in which case the soil scientists should be equipped with some quantitative and geospatial skills, while the data scientists should develop sufficient understanding of pedological principles (Arrouays et al., 2017a; Bui et al., 2020; Grunwald et al., 2011; MacMillan, 2008; Nenkam et al., 2023; Paterson et al., 2015). DSM demands substantial computational power for modelling large areas at high spatial resolution, which can pose problems for underfunded national institutions.

Recent developments point towards expanding data sources and applications. The increasing availability of proximal sensing, geophysical data, and citizen science observations is broadening the data base for DSM (McBratney et al., 2025; Shepherd et al., 2022; Stenberg et al., 2010; Terhoeven-Urselmans et al., 2010; Viscarra Rossel et al., 2016).

Advances in DSM are increasingly focusing on translating model outputs into accessible, web-based platforms and decision-support tools. WebGIS applications, cloud-based analytical platforms, and interoperable soil information systems now enable non-specialist users to explore, visualise, and apply soil information for land management and policy decisions, thereby bridging the gap between soil data production and practical use (e.g. [Aroca-Fernandez et al., 2025](#)).

There is also a growing shift from mapping individual soil properties towards digital soil assessment (DSA), which translates DSM outputs into indicators of soil functions, land capability, and crop suitability (Malone et al., 2017). DSA frameworks integrate multiple soil properties, thresholds, and risk indicators into actionable products such as erosion risk and crop suitability maps.

An ongoing, broader methodological debate concerns the relative roles of CSM and DSM. There is increasing recognition that data-driven modelling alone is insufficient and that insight into pedological processes should be embedded into ML workflows to ensure scientific plausibility, interpretability, and explainability (Wadoux et al., 2020).

Overall, the soil information field is moving toward a more integrated paradigm in which soil mapping, data management, and modelling are combined within unified information systems. This recognises that reliable soil information depends not only on advanced modelling techniques, but also on the integration of diverse data sources, the preservation of legacy knowledge, and the alignment of outputs with user needs. These developments provide the conceptual and methodological context for this study.

4. METHODS

4.1 Overview

This chapter describes the methods used to operationalise the theoretical frameworks of Chapter 3. It details the procedures applied to conceptualise the Namibian Soil Information System (NamSIS), identify, rescue, and revitalise Namibia’s legacy soil information and compile the Namibian Soil Profile Database (NSPD), compile a conventional national soil map of Namibia (NSMN), and develop national-scale predictive soil property maps. The methodological workflow is shown in Figure 10.

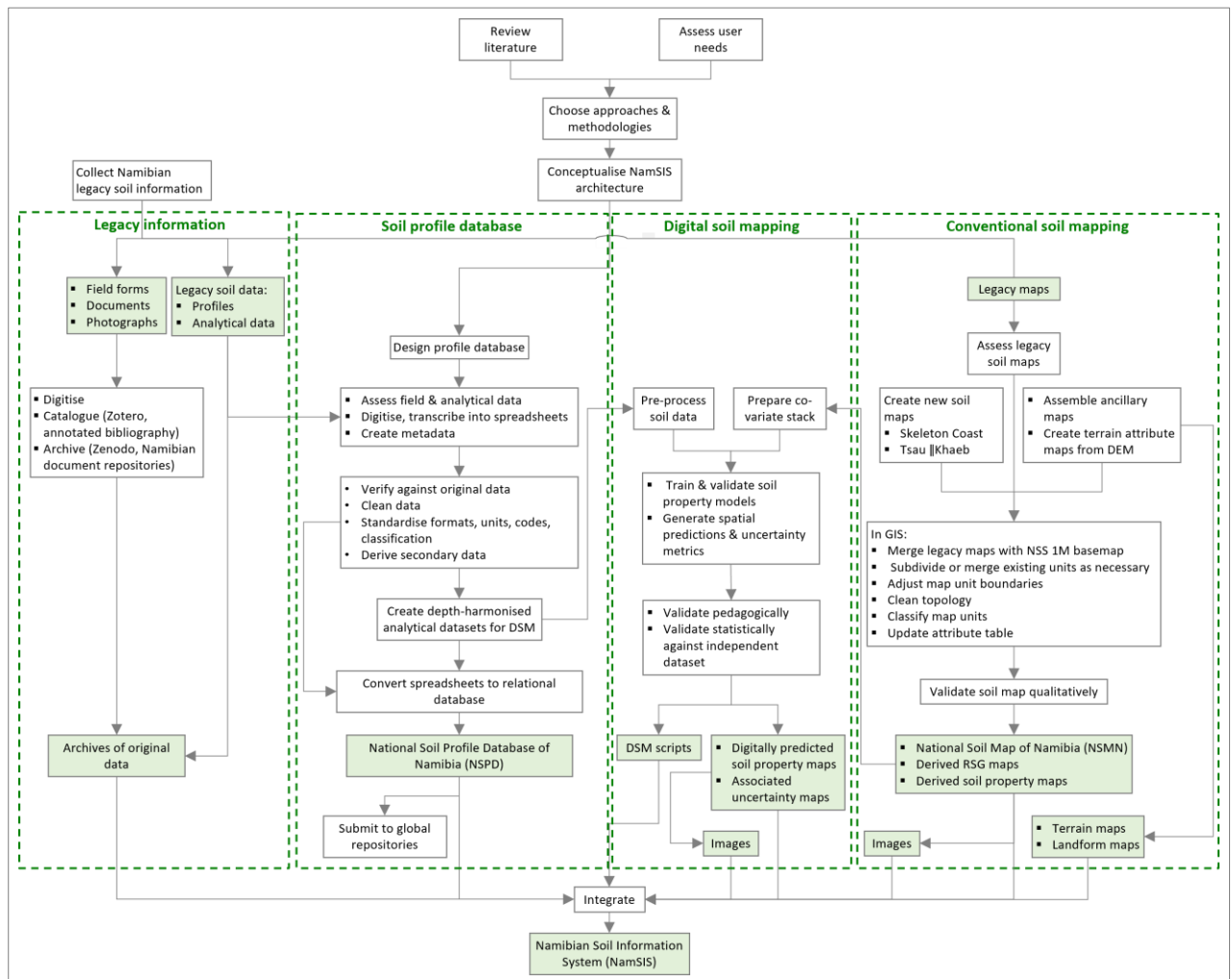


Figure 10. Workflow of the study — Development of a Namibian soil information system, national soil map, and spatial predictions of soil properties

The methodology directly addresses the three subsidiary hypotheses introduced in Chapter 2. Data discovery, digitisation, harmonisation, metadata compilation, and database design relate to the data integration hypothesis (H_1). The integration of harmonised legacy data with environmental covariates and expert knowledge informs the pedogenic–environmental relationship hypothesis (H_2), while the combined use of conventional soil mapping and machine learning addresses the conventional–digital convergence hypothesis (H_3).

4.2 Conceptualisation and design of the Namibian Soil Information System (NamSIS)

The conceptual design of NamSIS emerged from the recognition that a national soil information system must simultaneously serve diverse user groups and be technically sound, extendable, and compatible with international standards. Rather than functioning as a stand-alone database, NamSIS was conceived as a modular yet integrated digital platform supporting legacy data stewardship, data curation, conventional soil mapping, and digital soil predictions.

Design decisions were informed by the challenges and opportunities identified in Chapter 3, particularly the fragmented nature of Namibia's soil information, no prospect for national resurvey in the near future, and the need for alignment with global soil information initiatives and FAIR principles. The conceptualisation process involved three main steps: assessment of national user requirements, review of international soil information systems and standards, and development of a modular system architecture capable of incremental expansion.

This study primarily focussed on soil science and thus stopped short of developing fully interactive web or mobile user interfaces.

4.2.1 Assessment of user needs

User requirements were identified through workshop deliberations conducted during recent research projects³ and through long-term engagement with individuals requesting soil information from the author. These consultations revealed a diverse user community including farmers, agricultural researchers and extension officers, spatial and development planners, environmental scientists, engineers, educators, investors, and policy makers.

Information needs ranged from detailed soil profile descriptions and laboratory data required for research applications to generalised soil maps and derived soil function indicators used in land-use planning, environmental assessment, and climate adaptation. Frequent use cases included soil fertility management, crop suitability assessment, land degradation monitoring, carbon dynamics modelling, conservation planning, infrastructure site evaluation, and estimation of production potential for land reform initiatives. Incorporating these user needs ensured that system design was based on practical demand rather than purely technical considerations.

4.2.2 Existing systems and standards

The conceptual design drew on international soil information models and data standards, including WoSIS (Batjes et al., 2017, 2020, 2024; Batjes & Ribeiro, 2019; Calisto et al., 2023; Carvalho Ribeiro et al., 2020), AfSIS (ISRIC-WSL, 2023; Leenaars et al., 2014), GlobalSoilMap (Arrouays et al., 2014, 2017b, 2020b, 2020c), SOTER (ISRIC, 2019; van Engelen & Dijkshoorn, 2013; van Engelen & Wen, 1995), SoilGrids (Hengl et al., 2014, 2015, 2017a, 2017b), the Harmonized World Soil Database (FAO & IIASA, 2023; FAO/IIASA/ISRIC/ISSCAS/JRC, 2012), the WRB (IUSS

³ Agro-ecological zoning and carrying capacity assessments; Development of an integrated management plan for the Skeleton Coast–Iona National Park Transfrontier Conservation Area (SCIONA TFCA); Bush encroachment and control assessments; and Spatial biodiversity assessments.

Working Group WRB, 2015, 2022), and the FAO Guidelines for Soil Profile Description (FAO, 1990, 2006). The review provided a reference for aligning NamSIS with global infrastructures while taking into account national requirements.

4.2.3 System architecture concept

NamSIS was envisaged as a dynamic, modular system centred on a relational database, the Namibian Soil Profile Database (NSPD), to store soil profile-, horizon-, laboratory-, and metadata. The database structure accommodates both raw and depth-harmonised observations, for transparency and to enable future reprocessing as standards evolve.

Surrounding this core, the system architecture was designed to facilitate data access, spatial visualisation, and analytical procedures. The modular structure can support incremental growth, allowing new datasets, data types (e.g. proximal sensing; soil biological and microbiological data), models, and visualisation tools to be integrated without restructuring the core database.

4.2.4 Alignment with FAIR principles

Usability, standardisation, transparency, reproducibility, and long-term stability were central design principles. The system was intended to be accessible to non-specialist users while still providing sufficient analytical depth for advanced scientific applications.

To enable future data exchange with platforms such as WoSIS and GloSIS, emphasis was placed on format standardisation, including metadata reporting, database structures, attribute naming conventions, coding systems, depth harmonisation, uncertainty representation, and version control.

In line with the conceptualisation of soil survey as a knowledge system (Bui, 2004), NamSIS is designed not solely as a data repository, but as a unified system for capturing, harmonising, and sustaining soil knowledge across legacy and newly derived datasets. It is envisaged as a long-term national asset rather than a project-specific deliverable. Design choices therefore looked at durability, data stewardship, and institutional relevance. The system must support diverse land management, national planning, research, and reporting needs but must also remain adaptable to future scientific and technological developments.

4.3 Legacy data collection, rescue and revitalisation

The second objective addressed the risk of irreversible data loss associated with uncatalogued, ageing analogue materials, and institutional memory loss. The work proceeded through three interrelated stages: *data archaeology*, *data rescue*, and *data renewal*.

4.3.1 Collection and assessment of legacy materials (data archaeology)

A comprehensive search was conducted to locate, retrieve, document, and assess Namibia's scattered legacy soil information. Many of these materials had never been digitised or catalogued, were half-forgotten in filing cabinets and storerooms, and might have been lost permanently as

staff retire or leave institutions.

Materials were collected from ministries, parastatals, universities, research institutes, engineering firms, and individual researchers and consultants involved in historical soil surveys, and agricultural and environmental projects. Reconnaissance and irrigation-oriented studies from the 1950s to 1980s were retrieved from South African institutions, while published articles, theses and dissertations were obtained from libraries and online.

The search unearthed journal articles, dissertations, technical reports, paper maps, profile descriptions, hand-written field forms and notes, laboratory records, photographs, and both digital and analogue project-specific datasets. Each item was assessed for provenance, spatial coverage, scale, observational and analytical detail, and potential relevance to national soil information needs. The assessment provided context for decisions on data inclusion and prioritisation.

4.3.2 Digitisation and documenting (data rescue)

Collected materials were converted to digital formats to prevent further loss. By preserving the original paper records as scanned documents, they remain available for future re-examination.

Analogue maps were digitised manually or scanned and vectorised, and stored as GeoTIFF or shapefiles. Maps were georeferenced in QGIS (QGIS Development Team, 2024a) using control features such as graticules, road intersections, river confluences, and other landmarks.

Documents relevant to soil mapping, classification and characterisation in Namibia were gathered in a collection of PDF files. Some were already available in digital format, while hardcopies were scanned. These were archived in two national repositories, for redundancy and long-term accessibility. The documents were catalogued using reference-management software (Zotero), and an annotated bibliography was compiled and published on Zenodo.

Profile description field forms and laboratory records were scanned to PDF format. Profile, horizon and analytical data were transcribed into spreadsheets, which would eventually evolve into the Namibian Soil Profile Database. During this process, inconsistencies and obvious mistakes in both field forms and old spreadsheets were investigated and corrected, where possible.

Site and profile photographs were saved in various image formats (PNG, JPEG, TIFF).

Metadata were compiled for profile description and analytical datasets to document their nature, origin, data ownership, collection date, coordinate reference system, analytical methods, spatial distribution, links to photos, published articles or grey literature, and contact details of organisations or individuals.

4.3.3 Development of the Namibian Soil Profile Database (data renewal)

The renewal phase transformed rescued legacy data into the Namibian Soil Profile Database (NSPD). It integrated soil profile, augering, and sampling data from multiple sources into a harmonised structure for data curation, forming the core of NamSIS and providing the empirical basis for both conventional and digital soil mapping.

4.3.3.1 Source data

The principal datasets incorporated into the NSPD are outlined below and shown in Figure 10. Dataset acronyms have been retained from the original surveys and generally refer to the relevant project, sampling campaign, or geographic region. Further details on the origin, scope, and characteristics of each dataset are provided in the metadata of the soil profile database.

NARIS soil dataset

The *first National Soil Survey of Namibia (NSS)* was conducted between 1998 and 2000 under the Agro-ecological Zoning (AEZ) Programme of the Ministry of Agriculture⁴ in collaboration with the Cartographic Institute of Catalonia (ICC & MAWRD, 2000). It generated a substantial body of profile and horizon observations. Field observations followed the FAO Guidelines for Soil Profile Description (FAO, 1990) and were recorded by hand on paper forms, and much of the information was entered manually into four separate MS Access relational databases. Two of these databases (NSD1000 and NSD250) followed a SOTER-style structure (van Engelen & Wen, 1995), with tables covering terrain and terrain components, soil components, profile and horizon descriptions, laboratory results and methods, climate stations and climate data, map sources, data sources, and code lists. The remaining two databases (NSDkava and NSDowa) used a simpler structure organised around map units and map components, profiles, horizons and analytical tables, supported by code tables and data-entry forms.

The Ministry of Agriculture's Agro-ecological Zoning (MAWF–AEZ) team consolidated the profile, horizon, and laboratory analyses tables from NSD1000, NSD250, NSDkava, and NSDow — renamed NAM NSS, TWO NSS, OKA NSS, and OWA NSS — in a single MS Excel file. Additional *MAWF–AEZ survey campaigns* conducted between 2000 and 2009 contributed several datasets: AO, DOR, FEB, JAN, MAR, MAWF–AEZ MISC, NDONGA, OMAH_A, OMAH_B, ONGO, OVI, SPERR, TSUM, and Uitkomst. The combined data were disseminated on CD through the Namibian Agricultural Resources Information System (NARIS). This *NARIS Soil Dataset* was maintained and updated until the end of 2009, by which time it contained 2 155 georeferenced records, although the level of profile description and analytical completeness varied greatly across entries.

All laboratory analyses associated with these data were performed by the Ministry's Agriculture Laboratory (AgricLab). Although the AgricLab was not ISO-accredited, laboratory procedures followed ISO-aligned quality management practices. Quality control relied on documented standard operating procedures, the routine use of replicates and internal standards, certified reagents, regular instrument maintenance and calibration, and participation in the Agricultural Laboratories Association of Southern Africa (AgriLASA) inter-laboratory proficiency testing scheme.

During the present study, the profile, horizon, and laboratory tables of the NARIS Soil Dataset were ingested into the NSPD. Original NSS and MAWF–AEZ field forms and laboratory reports were re-examined, and additional field and analytical information not previously captured digitally

⁴ *Ministry of Agriculture* is primarily used in this dissertation, as the name and departments of the Ministry changed over the years, but the agricultural mandate remained constant: Ministry of Agriculture, Water and Rural Development (MAWRD; 1991–2000); Ministry of Agriculture, Water and Forestry (MAWF; 2000–2020); Ministry of Agriculture, Water and Land Reform (MAWLR; 2020–2025); Ministry of Agriculture, Water, Fisheries and Land Reform (MAWFLR, since 2025).

was extracted and transcribed. Compared with other project datasets, the NARIS Soil Dataset provided the largest number of observations and therefore formed the structural backbone of the NSPD.

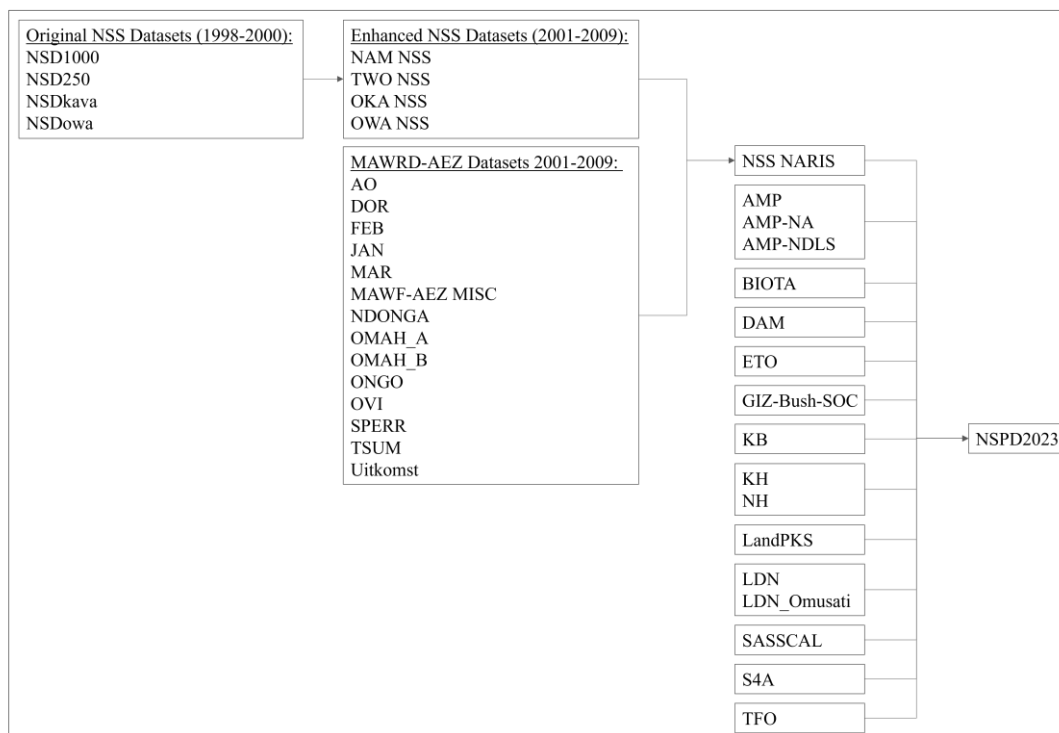


Figure 11. Datasets contributing to the Namibian Soil Profile Database.

AMP datasets

The *African Soil Microbial Genomics Project* (Cowan et al., 2022; Wild, 2016) contributed 179 topsoil (0–20 cm) samples across three datasets: AMP, AMP–NA, and AMP–NDLS. Eleven physicochemical soil parameters were analysed by the University of Pretoria following AgriLASA protocols (AgriLASA, 2004).

BIOTA dataset

The *Biodiversity Monitoring Transect Analysis in Africa (BIOTA) Project* established a continental network to monitor land degradation, support sustainable land use, biodiversity conservation, and rural development, and contribute to the Global Earth Observation System of Systems (GEOSS) (Petersen, 2008). The 11 BIOTA Observatories in Namibia contributed 638 samples from 257 profiles at three depths (0–10 cm, 10–30 cm, 30–60 cm), with analytical data for texture, pH (CaCl₂), electrical conductivity, total nitrogen, and organic carbon, as well as landscape and soil profile photos. Analyses were conducted at the University of Hamburg’s Institute of Soil Science.

GIZ–Bush–SOC dataset

The *Assessment of the Impact of Bush Encroachment and Bush Control on Soil Organic Carbon in Namibia Project* evaluated the effects of bush thinning on soil organic carbon and soil health under differing bush control and rangeland management strategies (Strohbach et al., 2023). The project contributed 162 topsoil samples (0–30 cm), analysed by the AgricLab for texture, pH (H₂O), bulk density, electrical conductivity, extractable calcium, magnesium, potassium and sodium, plant-available phosphorus, and organic carbon.

LandPKS

The *Land Potential Knowledge System (LandPKS) Project* developed a mobile-phone-based system for collecting and interpreting land potential data to support sustainable land management (Herrick et al., 2016). For the NSPD, the project contributed 918 point observations including slope, land use, land cover, vegetation cover, presence/absence of surface water, salt and cracks, effective depth, stoniness, and field estimated soil texture at seven pre-defined depths (0–1 cm, 1–10 cm, 10–20 cm, 20–50 cm, 50–70 cm, 70–100 cm, 100–150 cm). No laboratory data were generated, and some observations were collected by non-specialists. Consequently, the dataset required additional quality screening.

LDN, LDN–Omusati

The *Land Degradation Neutrality (LDN) Project* (Nijbroek et al., 2018), led by the Ministry of Environment and Tourism, contributed 319 topsoil (0–30 cm) and 277 subsoil samples (30–100 cm), including analytical data on organic carbon, bulk density, and particle size distribution. Analyses were performed by the AgricLab.

SASSCAL

The *Southern African Science Service Centre for Climate Change and Adaptive Land Management (SASSCAL)* Phase I subproject Monitoring agricultural ecosystems with regards to climate change effects (De Blécourt et al., 2018) contributed 181 samples from 30 profiles across four localities. Analyses by the University of Hamburg’s Institute of Soil Science included texture, bulk density, pH (H₂O and CaCl₂), electrical conductivity, total nitrogen, organic carbon, cation exchange capacity, and porosity. The dataset also included landscape and soil profile photographs.

TFO

The *Future Okavango (TFO) Project* aimed to improve knowledge-based land use management within the Okavango catchment under scenarios of global change (Pröpper et al., 2015). The project contributed 484 samples from 141 profiles, with analytical data for texture, pH (H₂O and CaCl₂), bulk density, electrical conductivity, cation exchange capacity, total nitrogen, and organic carbon, as well as landscape and soil profile photographs. Analyses were conducted at the University of Hamburg’s Institute of Soil Science and the University of Botswana’s Okavango Research Institute.

S4A

The EU-funded Horizon2020 *Soils4Africa (S4A) Project* contributed 194 samples from 112 locations in 28 clusters. Analyses were carried out by the South African Agricultural Research Council’s Institute for Soil, Climate and Water (ARC–ISCW): 63 samples using wet chemistry and 131 using mid-infrared spectroscopy (Paterson, 2021). Bulk density measurements for undisturbed surface, topsoil and subsoil samples from 10 locations were conducted by the AgricLab.

DAM, ETO, KH, NH, KB

Smaller datasets were from Etosha National Park (ETO) (Buch, 1990; Trippner, 1998), Khomas Hochland (KH) and Neudamm Hochland (NH) (Kempf, 1999), Damaraland (DAM) (Simmonds, 1994; Simmonds & Forbes Irving, 1995), and farm Klein Boesman (KB) (Strohbach et al., 2019).

The datasets varied widely in sampling design, analytical depth, and intended purpose, which

necessitated harmonisation and quality-control procedures.

4.3.3.2 Data entry, cleaning and quality control

Data entry and verification

The NARIS Soil Dataset served as the underlying structure for integrating additional datasets, gradually evolving into the NSPD. Data were managed within a master MS Excel workbook comprising four primary worksheets: *Profile Registration*, *Site Description*, *Horizon Field Data*, and *Horizon Laboratory Data*. The Profile Identification (PRID) functioned as the primary key, while the Horizon Number (HONU) linked individual horizons to their respective profiles.

Digital datasets were imported and aligned with the corresponding database fields. Paper-based records, including a substantial volume of previously uncaptured information from NSS and MAWF–AEZ field forms, were manually entered. During this process, existing digital records were systematically validated against original field forms, laboratory reports, and relevant grey literature. A number of MAWF–AEZ field forms could not be located and verified for transcription accuracy, although they were assessed against expert knowledge of those areas and comparable sites. These records were retained but flagged for future verification should the original documentation become available.

Transcription and georeferencing errors were identified and corrected, including coordinates located outside Namibia, omitted negative latitude values (southern hemisphere), and swapped latitude–longitude pairs. Where coordinates were absent or ambiguous, site locations could in some instances be inferred from textual descriptions of access routes or nearby landscape features. Duplicate profiles were identified and manually merged following verification.

Harmonisation and standardisation

Units of measurement were standardised (e.g. conversion from g kg^{-1} to %; meq kg^{-1} to cmol kg^{-1}). Site, profile, and horizon attributes were harmonised according to the FAO Guidelines for Soil Profile Description (2006) and SOTER conventions (ISRIC, 2019; Van Engelen & Dijkshoorn, 2013). Coordinates originally recorded in degree–minute–second format were converted to decimal degrees, and coordinate reference systems were documented.

Attribute definitions were verified to ensure consistency across datasets (e.g. distinguishing organic carbon from organic matter; confirming particle-size class definitions such as minimum sand fraction diameter). Soil classifications were translated, where feasible, from the South African Binomial system (MacVicar et al., 1977; Soil Classification Working Group, 1991), FAO/UNESCO (1974) and FAO/UNESCO/ISRIC (1990), Soil Taxonomy (Soil Survey Staff, 1999, 2014), and earlier version of WRB (ISSS/ISRIC/FAO, 1998) to the 2015 World Reference Base for Soil Resources (IUSS Working Group WRB, 2015). Reclassification was guided by diagnostic horizons, materials, and properties, profile descriptions and analytical data.

Quality control procedures

Quality control followed the principles outlined by Leenaars et al. (2014). Logical and numerical consistency checks were applied to identify errors and anomalies, for example:

- Valid value ranges (e.g. base saturation 0–100%);
- Particle-size fractions summing to $100 \pm 1\%$ (allowing for rounding);
- Sand fractions summing to total sand content;

- Horizon thickness corresponding to the difference between upper and lower depth limits;
- Consistency of classifications and properties (e.g. clay content <10% for Arenosols).

Many checks were conducted semi-automatically using spreadsheet filtering and formula-based validation tools. However, a substantial portion of data cleaning required expert-based evaluation. Automatically flagged values that were unrealistically high or low were examined individually, considering soil chemistry and physics, location, and pedological context.

Unverifiable (e.g. those with missing field forms) but internally consistent records were retained. Where inaccuracies, inconsistencies, or anomalies could not be resolved through logical correction or corroboration with original field forms or laboratory reports, the affected records were excluded from the database.

The data entry, standardisation and cleaning procedures transformed the heterogeneous legacy data into a structurally and analytically consistent dataset suitable for depth harmonisation and modelling.

4.3.3.3 Data enhancement and integration

Transfer functions for harmonising different phosphorus and organic carbon analytical methods could not be derived due to insufficient overlapping samples. These were, thus, kept in separate data fields.

Most profiles were sampled by genetic horizon, at various depths. For modelling purposes, data from standard depths are required. Thus, all points with data from at least three depths were depth-harmonised using an equal-area quadratic spline function (*ea_spline*, *ithir* package), fitted from the surface to a depth of 200 cm. Following some preliminary testing, a smoothing parameter (λ) of 0.1 was chosen to maintain acceptable fidelity to the observational data, while minimising over-smoothing. The data were aggregated into two sets of six and three depth classes, respectively. Increasingly, DSM is carried out at six depth intervals to conform with the GlobalSoilMap standard: 0–5, 5–15, 15–30, 30–60, 60–100, and 100–200 cm (Arrouays et al., 2017b). For most agricultural applications in Namibia, three depth classes are preferred by farmers and agricultural extension staff to represent topsoil, upper subsoil, and lower subsoil: 0–30, 30–60, and 60–100 cm.

Following harmonisation and quality control in the master MS Excel workbook, cleaned datasets were migrated into the relational database structure of the NSPD, which offers certain advantages over multiple standalone spreadsheets. It enforces data integrity through defined relationships between tables, reduces duplication and inconsistencies, and supports efficient querying, filtering, and aggregation of data. By separating data into discrete but interlinked tables (e.g. site, horizon, analytical data, metadata), the system allows complex queries and analyses that would be cumbersome and error-prone in flat spreadsheet structures. This modular architecture is scalable and adaptable, allowing individual components to be updated, expanded, or replaced without affecting the entire system, and facilitating integration of new datasets and variables.

The database design intentionally embeds a degree of redundancy, including the repetition of data in multiple formats (e.g. individual values, class ranges, class codes, and descriptive text) and the use of multiple measuring units. For example, dates of profile description are recorded in formats such as YYYY-MM-DD, DD-Mon-YYYY, as well as in separate Year, Month and Day fields, while organic carbon is recorded in both % and g/kg units. In addition, descriptive field names are

use alongside coded values. This approach facilitates integration with a range of international soil information systems, each of which may use different formats and coding conventions, while also improving accessibility for users unfamiliar with specialised coding schemes. Although such redundancy increases storage requirements, it enhances usability, flexibility, and compatibility across platforms and user groups.

Each point observation is uniquely identified by a primary key (PRID; profile identification), which links all associated site, horizon, and analytical data. Horizon-level observations are further distinguished by the secondary identifier HONU (horizon number), allowing structural representation of vertical soil profile information within the relational framework.

The standardised, quality-controlled original dataset, based primarily on genetic horizons, is retained alongside the two depth-harmonised datasets (with three and six depth classes respectively). The harmonised datasets offer data for DSM and direct integration with external systems, while the original dataset enhances transparency, reproducibility, long-term scientific value, and a safeguard against information loss during data transformation.

The soil profile database will be maintained as two versions. A public version for dissemination will ensure accessibility. The authoritative master version will be used for data curation, quality control, updates, and advanced analysis, and will protect data integrity. This separation will reduce the risk of unintended modifications to the core dataset. It will also allow incomplete or sensitive data to be managed appropriately.

4.4 Production of a conventional national soil map of Namibia

Namibian soil information consisted of miscellaneous survey products generated for different purposes and at different times and scales. Although they represented substantial expert knowledge, their analogue formats, incomplete coverage, disparate scales, and inconsistent typologies restricted their direct use in GIS-based planning and modelling. Historically, surveyors used manual methods to overlay information and delineate map units on hard copies of topographic maps, aerial photographs, or satellite images. While adhering to conventional soil mapping philosophy, this study integrated existing survey materials with newly captured data and supplementary environmental datasets within a digital geospatial environment to generate a harmonised, WRB-aligned national soil map. The following sections describe the primary source datasets and ancillary information used, and the procedures applied for data integration and soil map unit classification.

4.4.1 The basemap

The National Soil Map of Namibia (NSMN) built on the baseline created during Namibia's first National Soil Survey (NSS) (ICC & MAWRD, 2000) — see Section 4.3.3.1. Field observations were originally codified using the FAO Guidelines for Soil Profile Description (FAO, 1990) and soil classification was carried out using the 1998 edition of the World Reference Base for Soil Resources (ISSS/ISRIC/FAO, 1998).

The NSS produced four pedo-morphological soil maps, each accompanied by a relational database, covering different regions of the country at varying levels of spatial detail (Figure 12):

- *NSS 1M*: A preliminary national soil map at a scale of 1:1 million, consisting of 321 soil map units and 89 soil–terrain classes. This map excluded Etosha National Park, Skeleton Coast National Park, and the Sperrgebiet.
- *NSS TWO*: A 1:250 000 scale soil map covering 141 550 km² of Growing Period Zones 2 and 3 in the northeast of Namibia, where comparatively higher rainfall supports greater agricultural productivity.
- *NSS OKA*: A 1:100 000 scale soil map of the Okavango River Terrace, an area characterised by relatively high population density and subsistence-level rainfed smallholder agriculture plus a handful of state-owned irrigation schemes. The map consists of 198 soil map units and 17 soil–terrain classes, covering 13 183 km².
- *NSS OWA*: A 1:100 000 scale soil map of the central Cuvelai Basin, encompassing 87 645.5 km² across the Omusati, Oshana, Ohangwena, and Oshikoto Regions. This area is characterised by high population density and subsistence-level rainfed smallholder agriculture, and the map consists of 300 soil map units and seven soil–terrain classes.

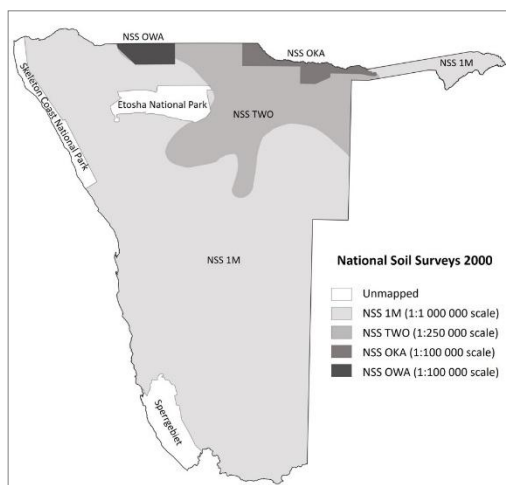


Figure 12. Locations of the four NSS surveys (ICC & MAWRD, 2000).

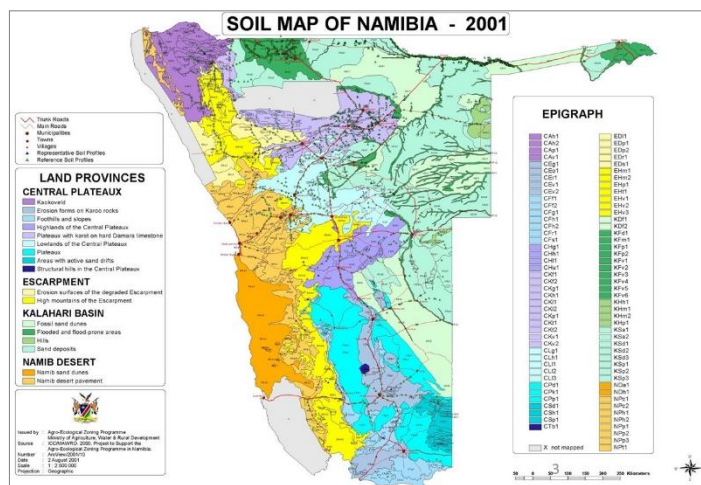


Figure 13. Preliminary Soil Map of Namibia (Coetzee, 2001).

The NSS 1M and NSS TWO datasets were subsequently transformed into a digital land resources information system, the *Namibian Soil and Terrain Digital Database (NamSOTER)* (Coetzee, 2001), following the SOTER methodology of van Engelen & Wen (1995). This process represented the spatial extent and topology of soil–terrain units as polygons within a GIS, linked to attribute information stored in a relational database management system. NamSOTER adopted a hierarchical structure comprising 89 SOTER/Terrain Units, defined primarily by dominant slope, relief intensity, and general lithology; 118 Terrain Components, based on meso-relief, slope, surface drainage, and depth to bedrock; and 269 Soil Components, characterised by diagnostic soil properties. These units were supported by 56 representative soil profiles (van Engelen & Dijkshoorn, 2013). NamSOTER's spatial component became known as the *Preliminary Soil Map of Namibia* (Figure 13) and was the most widely used (although incomplete) soil map of the country until the present study. NamSOTER was later incorporated into the *Soil and Terrain Digital Database for Southern Africa (SOTERSAF version 1.0)* (Dijkshoorn, 2003; FAO/ISRIC, 2003), which harmonised soil and terrain data from eight southern African countries.

Together, the NSS datasets and their NamSOTER derivative formed the primary spatial and conceptual backbone for the construction and subsequent refinement of the NSMN.

Global and continental products (DSMW; HWSO; Soil Atlas of Africa; SoilGrids; AfsIS) were

assessed but not incorporated as they conveyed less information than the NSS products and other legacy maps.

4.4.2 Supplementary legacy soil maps and databases

Additional spatial and attribute information from a range of legacy soil maps and databases was incorporated during the mapping process to improve coverage, resolution, and pedological detail.

MAWF–AEZ surveys

Between 2000 and 2009, the Ministry’s AEZ Programme team surveyed under-sampled areas. The most significant of these field campaigns included: the Auas-Oanob Conservancy in the Khomas Region ([Kutuahupira & Mouton, 2006a](#)); eastern communal lands of the Omaheke Region ([Kutuahupira et al., 2003](#); [Strohbach et al., 2004](#); [Strohbach & Kutuahupira, 2014](#)); Uitkomst Research Station ([Coetzee et al., 1999](#)); Omahanene Research Station ([Kutuahupira et al., 2001a](#)); Sonop Research Station ([Kutuahupira et al., 2001b](#)); and a 100 km by 200 km block between Windhoek and Gobabis ([Coetzee, 2009](#); [Kutuahupira & Mouton, 2006b](#)). Most AEZ outputs existed only as working maps in project reports, laboratory analyses, and georeferenced site and profile descriptions on handwritten field forms, partially captured in the internal National Agricultural Resources Information System (NARIS) (MAWRD-AEZ, 2009). These datasets substantially enriched the NSMN.

Older (“pre-Independence”) soil surveys

Prior to Namibian independence in 1990, several South African governmental, academic, and private research institutions conducted soil surveys for the former South West Africa Administration (*sic*), particularly for the Department of Water Affairs and homeland administrations: [AOC Technical Services \(Pty\) Ltd., 1967a, 1967b, 1967c, 1966](#); [Department of Water Affairs, 1987](#); [Du Plessis and Dohse, 1978](#); [Ellis et al., 1980](#); [Lambrechts, 1977](#); [Lambrechts & Schloms, 1993](#); [Louw, 1957](#); [Loxton et al., 1974a, 1974b, 1970](#); [MAWRD, 1995](#); [Murray et al., 1997](#); [Opperman et al., 1982](#); [Page, 1980, 1979, 1978, 1976](#); [Partridge, 1987a, 1987b](#); [Van Woerkom & Louw, 1971](#). These studies primarily focused on assessing irrigation and rural development potential and ranged from reconnaissance to detailed mapping scales. The technical reports and resulting maps mostly remained unpublished and were obtained from various institutions and personal collections. They contain valuable information for soil unit characterisation. Some challenges were encountered: As the surveys predated GPS technology, most soil profile descriptions were not georeferenced, with locations described textually or indicated approximately on schematic maps. Different versions of the evolving South African Soil Classification System were used ([MacVicar et al., 1977](#); Soil Classification Working Group, 1991), complicating the correlation to the WRB 2015 system. In some cases, technical reports contained field observations and laboratory results, but the corresponding maps could not be located, and coordinates could therefore not be assigned to point data.

Contributions from environmental projects

Several environmental research projects generated observations that, while originally collected for ecological or biodiversity purposes, supplied useful soil information (see Section 4.3.3.1).

Extensive research on landscape ecological risk was undertaken in *Etosha National Park* during the 1990s under the DFG/GTZ Research Cooperation Project *Environmental Change in the Etosha*

National Park/Northern Namibia (Az. Bu 659/4-1) — also known as the *LERIS* Project. Led by Dr Manfred Buch, this work resulted in several diploma and doctoral dissertations, scientific publications, and internal reports (e.g. [Beugler-Bell, 1991](#); [Beugler-Bell et al., 1993a, 1993b](#); [Beugler-Bell & Buch, 1997](#); [Buch, 1997, 1993, 1990](#); [Buch et al., 1997](#); [Buch & Trippner, 1997](#); [Trippner, 1998, 1997](#)). These materials were instrumental in filling the mapping gap in Etosha left by the NSS.

Other sub-national datasets include work by Simmonds (Simmonds, 1998, 1994; Simmonds & Burke, 2000; Simmonds & Forbes Irving, 1995) in the *Kunene* and *Kavango* Regions, as well as *Environmental Profiles of Caprivi, North-Central Namibia*, and the *Kavango* by (Mendelsohn et al., 2000; Mendelsohn & El Obeid, 2003, 2004; Mendelsohn & Roberts, 1997).

Large multidisciplinary initiatives such as *BIOTA Southern Africa*, *SASSCAL*, and *The Future Okavango (TFO)* did not produce maps but contributed detailed soil profile descriptions, laboratory data, and photographs from ecological monitoring sites spanning major Namibian biomes. Additional laboratory data — primarily soil organic carbon and bulk density — were obtained from the *GIZ-Bush-SOC* and *Land Degradation Neutrality (LDN)* projects, while the *LandPKS* initiative provided field observations of soil texture, stoniness, topography, land use, and vegetation cover.

The *Atlas of Namibia* (Mendelsohn et al., 2002) included a generalised 1:4 million soil map derived from the NSS 1M. For areas not covered by the NSS, broad soil groups were inferred by Simmonds using the Land Systems of the Agro-ecological Zones map (De Pauw, 1996; De Pauw et al., 1999a, 1999b; De Pauw & Coetzee, 1999), the 1:1 million Geological Map of Namibia (Geological Survey of Namibia, 1980), the Etosha soil map of [Beugler-Bell & Buch \(1997\)](#), and visual interpretation of Landsat 5 imagery.

Although these datasets were highly diverse, they added valuable geographic coverage, particularly in regions underrepresented by the NSS products.

4.4.3 Ancillary information

Terrain attributes

Relief exerts strong control on soil depth, texture, drainage, erosion, and redistribution of materials across landscapes. Thus, terrain analysis provides spatial context for interpreting soil–landscape relations when soil profile data are scarce. For the construction of the national soil map, terrain attributes and secondary landform maps were not used as statistical inputs like in DSM, but as interpretive layers to guide delineation and refinement of map unit boundaries, and subdivision of heterogeneous units. These were derived from the 1 arc-second (~30 m) resolution Shuttle Radar Topography Mission (SRTM) DEM (Earth Resources Observation and Science (EROS) Center, 2000), using Geospatial Data Abstraction Library (GDAL) tools in QGIS ([GDAL Development Team, 2024](#); [QGIS Development Team, 2024a](#)). For applications with high computational demands, the 30 m data were resampled to 90 m.

Elevation — the height of a point on the earth’s surface, relative to mean sea level — can be used to characterise landforms (e.g. distinguish between uplands and lowlands), calculate evapotranspiration (IIASA & FAO, 2012; Monteith, 1965, 1981) and define thermal zones (IIASA & FAO, 2012).

Slope gradient — the inclination of the land surface relative to the horizontal (Schoeneberger et al., 2017) — fundamentally influences soil formation and distribution. It affects the kinetic energy, erosivity, infiltration rate, sediment-carrying capacity of surface runoff, and is inversely related to soil moisture retention time (Schoeneberger et al., 2017). It affects water flows, access to steep terrain, suitability for planting crops and constructing infrastructure, and the potential for landslides and flooding (Jenness, 2013). In general, steeper slopes are associated with thinner soils, higher erosion rates, and reduced agricultural suitability, while gentler slopes favour deeper profiles and greater soil development. Slope was mapped both as a continuous raster and as classes based on the Global Agro-Ecological Zones (GAEZ) categories (IIASA & FAO, 2012). The proportion of land within each soil map unit exceeding an 8% slope threshold — widely used as a practical threshold for mechanised cropping and an indicator of elevated erosion risk — was calculated.

Slope length and *form* were not considered in the national soil map, as these attributes are highly scale dependent and better suited to local erosion modelling and site-specific assessments.

Aspect — the downslope orientation expressed as an azimuth from true north — influences insolation, temperature regimes, evapotranspiration, vegetation patterns, and soil moisture dynamics. In arid and semi-arid environments, aspect-related differences have impacts on organic matter accumulation and biological activity, particularly on steeper slopes. Aspect was mapped both as a conventional categorical raster and as a *combined slope–aspect* visualisation, in which flat areas (0–2% slope) were masked and steeper terrain emphasised through progressive colour saturation. These visual products were particularly valuable for identifying subtle landscape features and asymmetric hillslope patterns that are not readily apparent from elevation or slope maps alone.

Expressions of *terrain complexity* provide both visual and quantitative insight into landscape dissection, erosion potential, and soil heterogeneity. Various metrics were considered, including roughness, degree of dissection, standard deviations of slope and altitude, curvature (plan, profile, tangential, and contour curvature), terrain ruggedness indices of Riley (Riley et al., 1999) and Nellemann (Nellemann & Fry, 1995), site ruggedness index, basin-scale ruggedness (drainage density), rugosity (2D/3D area ratio), Melton ruggedness number (Conrad, 2012), 1st derivative of flow direction, and Hobson’s surface roughness factor (Hobson, 1972). All these measures are highly sensitive to DEM resolution and neighbourhood analysis window (NAW) size. After exploratory testing, the *Riley Terrain Ruggedness Index (TRI)* (Riley et al., 1999) was selected for national-scale mapping. TRI quantifies the mean absolute difference between the central cell and each surrounding cell within the NAW (Wilson et al., 2007), offering an intuitive measure of local relief variability. TRI values were reclassified into six terrain ruggedness classes, and the proportion of each soil map unit falling within each TRI class was calculated using zonal histograms. TRI was used to refine topographic boundaries, such as distinguishing between valley floors and canyon walls, and identifying rock outcrops and inselbergs on plains. It also supported interpretation of soil depth variability, erosion susceptibility, and lithological control within map units, particularly in the escarpment and mountainous regions.

Relative or local relief — the elevation difference between the highest and lowest points within a defined area — captures landscape variability independently of absolute elevation (Schulze & Horan, 2007). Although no national map of relative relief was produced due to its strong sensitivity to NAW size, the concept informed interpretation of highlands, scarps and valley systems in selected regions.

The *Topographic Position Index (TPI)* (Jenness, 2002, 2006; Weiss, 2001), which expresses the relative elevation of a location compared to its surroundings, proved too generalised to be operationally useful at national scale, but was used selectively for local analysis. Positive TPI values indicate ridges or upper slopes, while negative values correspond to valleys and depressions.

Drainage pattern analysis revealed a range of network types — karstic, centripetal, radial, dendritic, annular, pinnate, trellis and parallel — that reflects variations in bedrock geology, structure and geomorphic evolution. The karstic pattern was particularly helpful to delineate and classify calcareous landscapes.

Hydrological attributes

Hydrological features derived from the DEM — including *flow direction*, *flow accumulation*, *catchments*, and *drainage networks* — supported interpretation of water movement and sediment redistribution. Drainage patterns — the recurring arrangement of interconnected drainage channels across the surface of a land area — reflect bedrock properties such as strength, fracture density, infiltration capacity, and how geological structures control the landscape (Schoeneberger et al., 2017), and they are particularly informative in arid environments where episodic runoff dominates geomorphic processes. The hydrological layers were useful for delineating alluvial soils, identifying zones of sediment accumulation, and refining boundaries along ephemeral river systems.

The *Topographic Wetness Index (TWI)* was explored as an indicator of potential water accumulation, with implications for soil depth, texture, salinity, and nutrient redistribution (Schillaci et al., 2015). It was not used directly in national-scale map unit definition but remains a useful tool at local scale to identify and interpret hydromorphic features in valley bottoms and floodplains.

Landform maps

Landform maps represent integrated expressions of relief, geology, and geomorphic processes, and thus offer a framework for understanding soil–landscape relationships. In the absence of a definitive Namibian terrain morphology map, various landform classification systems were evaluated and synthesised to support soil mapping.

The main contribution was a *Land Systems* map (Figure 14), compiled by the author during development of the updated Agro-Ecological Zones Map of Namibia (Van der Waal et al., 2021a, 2021c). It is a synthesis of Land Region information from the NSS 1M map, Land Units from the NSS OKA and NSS OWA maps, and supplementary data from the AEZ Land Systems (De Pauw, 1996; De Pauw et al., 1999a, 1999b; De Pauw & Coetzee, 1999), and NamSOTER (Coetzee, 2001). This composite landform representation provided a Namibia-specific geomorphic context that aligned well with observed soil patterns.

Additional global and continental landform products were consulted: the United States Geological Survey (USGS) Land Surface Forms (Cress et al., 2009; Sayre et al., 2014); the Earth Systems Research Institute's (ESRI's) Ecological Facets Landform Classes (ESRI, 2020; ESRI & USGS, 2015); and two European Soil Data Centre (ESDAC) global landform maps (ESDAC, 2008) based on the methodologies of Meybeck et al. (2001) and Iwahashi (Iwahashi et al., 2018; Iwahashi & Pike, 2007). Although these products vary in conceptual basis and class definitions, they were useful for cross-checking interpretations.

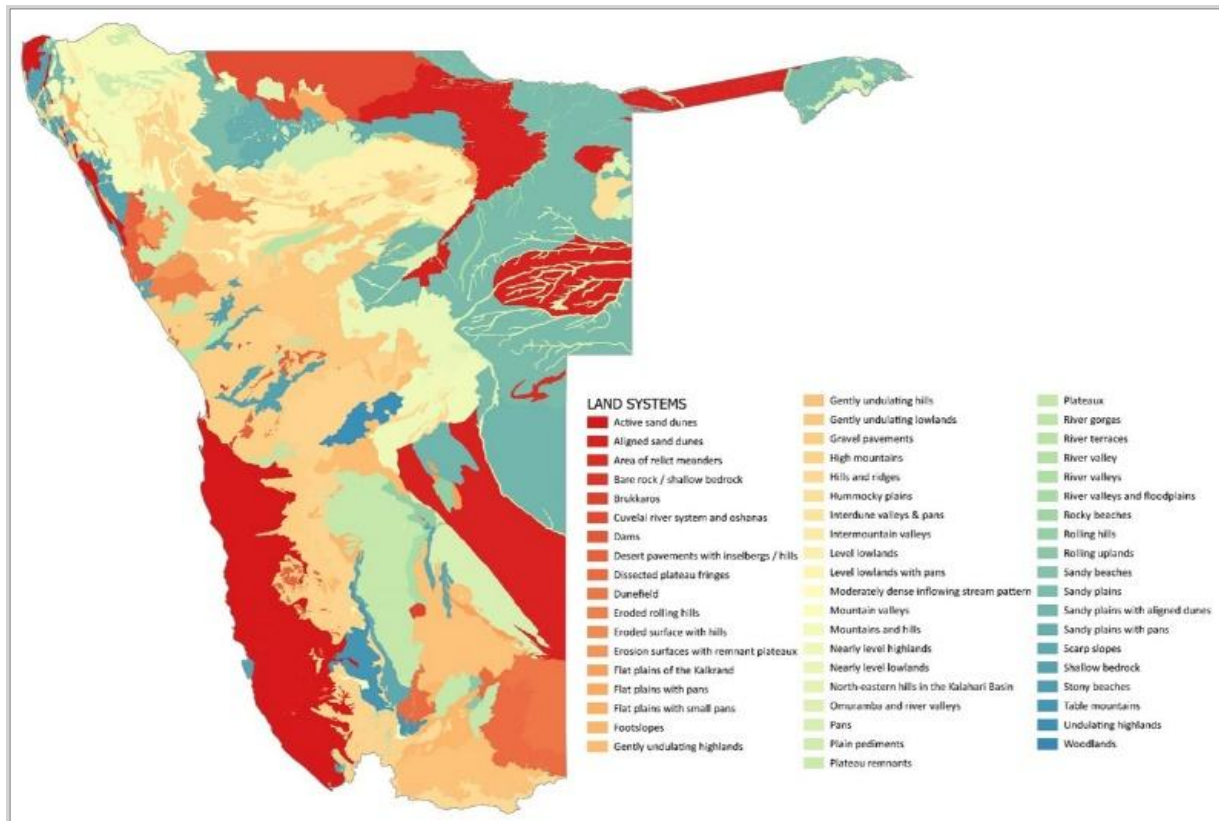


Figure 14. Land Systems of the 2021 Agro-Ecological Zones Map of Namibia (Van der Waal et al., 2021a, 2021c).

Geology

Soil formation is significantly influenced by the mineral composition and weathering behaviour of the parent material, thus the *digital 1:250 000 Geological Map of Namibia* (Geological Survey of Namibia, n.d.) guided the delineation of soil map units where lithological boundaries corresponded to transitions between soil types. Some inferences could be made about the mineralogy, texture and chemistry of soils, based on the weathering products of the rock types. Geology also influences topography and hydrology. For example, resistant lithologies like quartzite and dolerite form ridges with well-drained shallow soils, while easily weathered shales form deeper, finer-textured soils with greater moisture retention on footslopes and in valleys.

Remote sensing

Natural colour *satellite images* (Bing Aerial; Google Hybrid) were used as backdrops for online digitising and editing of features in QGIS to refine boundaries and validate existing polygons. Land surface colour, texture, and geomorphically distinct features (e.g. dune fields vs. gravel plains; clay-rich pans vs. sandy inter-dune depressions; calcrete plateaux dotted with dolines) were especially valuable in hyper-arid western Namibia, where field access is challenging and historic surveys were sparse. No spectral analysis was done on satellite images.

Soil profile data

Georeferenced profile descriptions and analytical data from the *Namibian Soil Profile Database* (see Section 4.3.3) served as ground control points. The database provided both qualitative and quantitative information to draw boundaries around clusters of profiles that share similar analytical and morphological characteristics, and to infer diagnostic horizons, properties and thresholds for

classification.

4.4.4 Completion of previously unmapped regions

Two large areas —*Skeleton Coast National Park* and *Tsau ǁKhaeb National Park* (formerly the “*Sperrgebiet*”, meaning “forbidden area”) — were mapped for the first time in any detail, using satellite imagery, terrain derivatives, geological information, and limited soil profile data. Map confidence is lower for the Tsau ǁKhaeb area, as access to the park is strictly controlled, and field observations by various botanists and geologists were consulted in the absence of dedicated soil surveys. These new soil maps resulted in full spatial coverage of Namibia at a scale larger than 1:4 million for the first time.

4.4.5 Integration — the National Soil Map of Namibia

The integration of the source maps produced at different scales posed a challenge, as they represented different levels of spatial generalisation and mapping intent. Small-scale maps prioritised national coherence and broad soil–landscape relationships, while larger-scale maps captured finer detail relevant to local land use and management. Preserving the finest detail across the country was neither practically feasible, nor would it be supported by data density or mapping confidence. The chosen strategy was to let the 1:1 million Preliminary Soil Map of Namibia (directly derived from the NSS 1M map) provide structural continuity at the national level, while downscaling it to a nominal 1:250 000 final scale with the aid of larger-scaled products from the NSS (1:100 000 and 1:250 000) and other legacy soil maps, the geology map (1:250 000), terrain attributes (30–90 m resolution), and satellite imagery. The information of source maps more detailed than 1:100 000 scale was generalised upward to ensure consistency with the national mapping scale.

Legacy soil maps were reviewed for original mapping purpose, extent, scale, spatial accuracy, completeness, typology, quality, and compatibility with the base map. Analogue maps were digitised or scanned, and all selected maps were imported into QGIS. They were georeferenced by aligning graticules and identifiable landmarks.

The NSS TWO, NSS OKA AND NSS OWA maps were merged with the Preliminary Soil Map through manual editing and edge-matching of polygons to ensure seamless transitions where maps met or overlapped. The map of Etosha was incorporated next. In all three instances, terrain and soil units had to be reassessed and correlated with those of the basemap. The newly mapped Skeleton Coast and Tsau ǁKhaeb/Sperrgebiet were made compatible from the outset. With national areal coverage achieved, spatial detail was refined through manual incorporation of the other source maps using the editing functions of QGIS.

Where necessary, map units were subdivided or merged to better reflect dominant soil patterns. Digital elevation models and derived terrain attributes, drainage information, satellite imagery, geology, vegetation, land systems, and agro-ecological zones were displayed as backdrops to the soil polygons and continuously consulted during the process. These layers supported decisions about the delineation, subdivision, merging, and adjustment of units. They were used as interpretive evidence and did not determine boundaries automatically.

On-screen manual boundary editing was typically carried out at working scales between 1:20 000

and 1:50 000, allowing sufficient spatial detail for feature differentiation while maintaining cartographic legibility. By digitising at large scales but expressly defining smaller intended usage scales, this approach avoids false precision that can arise when fine-scale detail is extrapolated beyond its evidentiary support. The final map can be used with high confidence for applications at 1:250 000 scale or smaller, and moderate confidence at scales between 1:100 000 and 1:250 000. Although the map may appear scale-independent within a GIS environment, positional accuracy at larger display scales is not supported by sufficient soil profile data, legacy maps or ancillary layers.

Following boundary refinement, polygon topology was cleaned up by removing gaps, overlaps and sliver polygons to ensure a consistent and valid geospatial dataset. The resulting map is designed to be revisable, allowing incremental updates and refinement as new soil observations become available.

No automated procedure was used to generate polygons or assign soil classes. In contrast to quantitative approaches based on automated terrain or parent-material delineation from digital elevation and remotely sensed data (for example, Dobos et al., 2005, 2010, 2013), the present compilation did not use automated procedures to generate polygons or assign soil classes. It was an expert-driven conventional soil-mapping exercise. Digital environmental layers supported but did not replace pedological interpretation.

4.4.6 Map unit design and classification

Soil classification and reclassification to the 2015 revision of the World Reference Base for Soil Resources (IUSS Working Group WRB, 2015) were also done manually. It was based on multiple lines of evidence, including soil profile descriptions, laboratory analytical thresholds (e.g. carbonate content, soil organic carbon, salinity), horizon morphology, and known pedogenic processes. The source maps employed several systems, including early versions of the South African (MacVicar et al. 1977; Soil Classification Working Group, 1991), FAO (FAO/UNESCO, 1974; FAO/UNESCO/ISRIC, 1990), and WRB (IUSS/ISRIC/FAO, 1998) classifications. Correlations presented by Fey (2010) were used where clear equivalences existed, but where direct correlation was not possible, the classification was reconstructed from textual soil-unit descriptions, profile descriptions, horizon morphology, and analytical data.

At the highest level of classification, Reference Soil Groups (RSGs) were identified using the WRB Key, and principal and supplementary qualifiers were assigned to express diagnostically relevant soil characteristics.

In this study, a soil typological unit (STU) represents a single, fully classified soil type, defined by its RSG and associated qualifier combination. STUs therefore correspond to the basic building blocks of the national soil map and represent conceptually homogeneous soil types as defined by the WRB classification system.

Each polygon in the National Soil Map of Namibia (NSMN) was assigned a single soil map unit (SMU) label. An SMU represents a landscape unit characterised by a recurring and interpretable pattern of soils rather than a single homogeneous soil body. To capture both dominant soil patterns and important secondary components, each SMU comprises from one to four STUs. The first-listed STU represents the dominant soil, estimated (not measured) to occupy more than 50% of the polygon area and to exert the strongest control on land-use potential and soil-landscape

processes. One additional STU may represent an associated soil, typically occupying approximately 25–50% of the polygon area and occurring in predictable landscape positions related to relief, parent material, or hydrological gradients. Where relevant, one or two further STUs represent included soils, generally occupying approximately 5–25% of the polygon area, which may influence ecological conditions or specific land-use or management decisions despite their small spatial extent. This hierarchical SMU–STU structure allows the national soil map to represent spatial soil heterogeneity while remaining legible and interpretable at national and regional scales, decreasing the unrealistic assumption of internal uniformity within map polygons

The map colour scheme, implemented through a QGIS *style* file, was based on the Soil Atlas of Africa (Jones et al., 2013), with some changes for better visual distinction between soil types, as shown in Figure 15.

Soil Atlas of Africa		New National Soil Map of Namibia
	Durisol	
	Fluvisol	
	Gleysol	
	Gypsisol	
	Regosol	
	Solonchak	
	Solonetz	
	Technosol	
	Vertisol	

Figure 15. Differences in the colour palettes of the Soil Atlas of Africa (Jones et al., 2013) and the new National Soil Map of Namibia

Metadata were created to document the location, provenance, and scale, of the contributing legacy sources. The classification framework provides a transparent and revisable basis for national-scale soil assessment and for integration with digital soil property modelling.

4.4.7 Quality control and validation

The draft national soil map and its classification framework were subjected to iterative consistency checks, field validation, comparison with legacy maps, and consultation with scientists familiar with Namibian soils and the NSS outputs. Validation focused on assessing the internal consistency of map units, the placement of map unit boundaries, soil transitions across these boundaries, and the correctness of WRB classifications. Special attention was paid to areas where legacy sources overlapped, where map units were generalised across scales, and where soil heterogeneity was high due to complex terrain.

Discrepancies or ambiguities identified during review were resolved through re-examination of legacy soil profiles, laboratory data, field photographs, terrain attributes, and geological context. Where multiple interpretations remained plausible, classification decisions favoured consistency with dominant soil-forming processes and regional soil patterns rather than local anomalies. The resulting map represents a consensus-based synthesis of available evidence rather than a definitive depiction of soil boundaries, and its confidence varies spatially in relation to data density, access constraints, and landscape complexity.

4.5 Predictive spatial modelling

4.5.1 Overview of DSM workflow

Digital soil mapping was undertaken to produce continuous spatial predictions of key soil properties across Namibia, to complement the categorical national soil map and fill long-standing gaps in quantitative soil information.

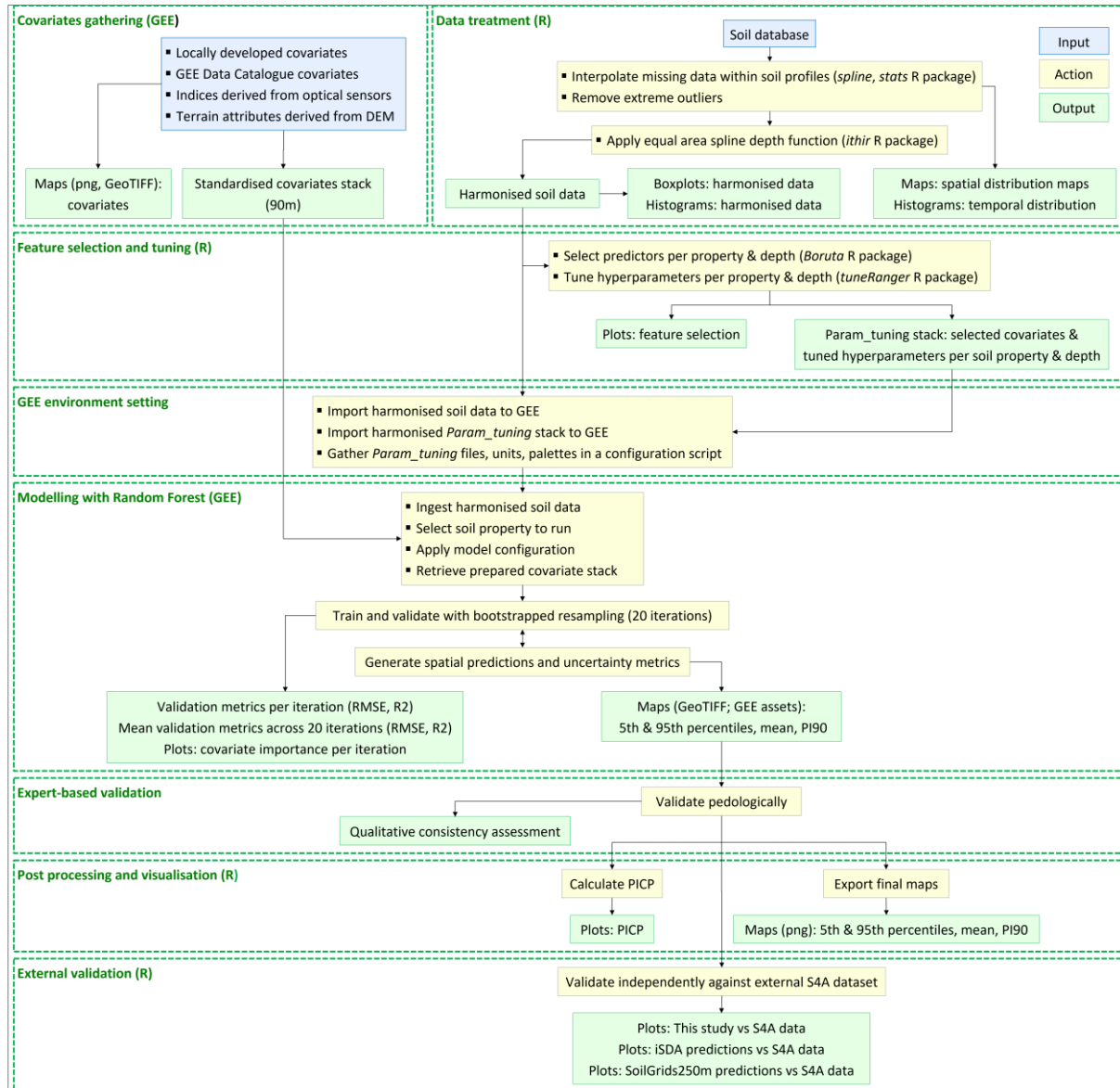


Figure 16. Digital soil mapping workflow.

Preparing for DSM required a harmonised geospatial workflow, shown in Figure 16. Soil profile data were subjected to quality control and depth-harmonisation to ensure comparability across disparate datasets and allow the models to operate on consistent numerical inputs. In parallel, a stack of environmental covariates was assembled to serve as predictors for modelling. A machine-learning algorithm was used to relate soil observations to the environmental covariates. The modelling process involved iterative training, feature (predictor) selection, hyperparameter tuning, and assessment of model performance, to ensure reliable predictions across Namibia's

heterogeneous landscapes. Bootstrap procedures were applied to spatially predict depth-specific soil properties at 90 m resolution across Namibia, and to generate uncertainty estimates corresponding to a 90% prediction interval. The final soil property maps were assessed for pedological plausibility and externally validated against an independent dataset.

4.5.2 Choice of modelling platform and algorithm

Google Earth Engine (GEE) was adopted as the primary modelling platform, mainly for its greater cloud-based processing power and direct access to environmental covariate datasets in the Earth Engine Data Catalogue. R was used for tasks that required finer user control but less computational power: data preparation, spline depth harmonisation, feature selection, hyperparameter tuning, and selected post-processing tasks such as map visualisation and independent validation. As GEE requires a subscription and imposes quotas on processing time, memory, and storage, some compromises had to be made regarding the spatial resolution and number of iterations.

A systematic comparison of multiple machine learning algorithms was beyond the scope of this study, which prioritised the development of an operational, reproducible, and updatable national-scale DSM framework. Random Forests (RF) was therefore selected as the main modelling algorithm after trials with alternatives yielded similar accuracy. RF is a proven, widely use, stable ML algorithm, and can be interpreted intuitively.

4.5.3 Soil data

The modelling of 16 soil properties (Table 1) was based on a harmonised legacy dataset comprising 4 958 georeferenced soil profiles and augerings and 13 362 individual samples from the Namibian Soil Profile Database that was compiled during this study (see Sections 4.3.3 and 5.1.2; [Coetzee, 2025](#)). Geographic coordinates were checked and corrected where necessary. To ensure positional reliability, only observations with a spatial precision better than 0.0001° were retained. Older datasets recorded in the Schwarzeck datum were transformed to WGS84. Attribute values were screened for extreme outliers or laboratory inconsistencies, and missing values within soil profiles were estimated through interpolation, using the *spline* function of the *stats* R package (R Core Team, 2024a, 2024b). Heterogeneous sampling depths required fitting of equal-area quadratic splines (*ea_spline* function of the *ithir* R package) to profiles with at least three depth observation, for depth normalisation prior to modelling. This approach follows established methods for depth harmonisation (T. F. A. Bishop et al., 1999; Malone et al., 2009; Ponce-Hernandez et al., 1986). A smoothing parameter ($\lambda = 0.1$) was selected to balance fidelity to observed measurements with the need to reduce abrupt changes between horizons. The resulting continuous depth functions were aggregated into three agriculturally relevant standard depth intervals (0–30 cm, 30–60 cm, and 60–100 cm) which formed the basis for all subsequent property-specific modelling.

Table 1. Soil properties selected for modelling

Soil property	Abbreviation	Unit
bulk density	BD	kg dm ⁻³
base saturation	BS	%
pH (1:2.5 soil:water)	pH (H ₂ O)	
cation exchange capacity	CEC	cmol(+) kg ⁻¹
extractable calcium, magnesium, potassium, and sodium	Ca, Mg, K, Na	mg kg ⁻¹
electrical conductivity (1:2.5 soil:water supernatant)	EC (1:2.5)	μS cm ⁻¹

electrical conductivity (saturated paste extract)	EC (sat)	$\mu\text{S cm}^{-1}$
organic carbon	OC	%
plant-available phosphorus	P	mg kg^{-1}
total nitrogen	N	mg kg^{-1}
sand ($>53 \mu\text{m}$), silt ($2\text{--}53 \mu\text{m}$), clay ($<2 \mu\text{m}$)		%

4.5.4 Environmental covariates

A total of 65 environmental covariates were assembled to represent the SCORPAN factors, incorporating both continuous and categorical variables.

Soil: The primary soil predictor was the National Soil Map of Namibia produced during this study (see Sections 4.4 and 5.2), classified according to the WRB Reference Soil Groups (IUSS Working Group WRB, 2015).

Climate: Climatic predictors included long-term temperature and precipitation variables from WorldClim Climatology v2.1 (1970–2000) (Fick & Hijmans, 2017), the Global Aridity Index and Potential Evapotranspiration datasets (Trabucco & Zomer, 2022; Zomer et al., 2022), and the ALOS Continuous Heat-Insolation Load Index (CHILI) (Theobald et al., 2015). These datasets capture moisture availability, evaporative demand, and thermal regimes relevant to soil formation.

Organisms: Spectral information, derived from Landsat 5 (1985–2011) and Landsat 8 (2013–2024), was used to represent soil and vegetation characteristics linked to texture, organic matter, mineralogy, and soil moisture (Mulder et al., 2011). Median surface reflectance values were calculated for two seasonal windows — winter (June–July) and summer (February–March) — using scenes with less than 1% cloud cover. Data from all 38 years were included to dampen inter- and intra-annual variability typical of arid-zone vegetation. Vegetation indices were derived within GEE: NDVI, EVI, SAVI, MSAVI, and kNDVI (Camps-Valls et al., 2021; Rondeaux et al., 1996; Wang et al., 2023). Additional biotic covariates included the ESA Global Land Cover Map (Sentinel-1/2; Zanaga et al., 2022) and Namibia-specific layers: vegetation types, average seasonal plant productivity (Atlas of Namibia Team, 2022), potential carrying capacity (Van der Waal et al., 2021a, 2021b), and agro-ecological zones (De Pauw, 1996; De Pauw et al., 1999a, 1999b; De Pauw & Coetzee, 1999; Van der Waal et al., 2021a, 2021c).

Relief: Topographic variables were derived primarily from NASADEM (30 m) (NASA JPL, 2020) and the ALOS DEM (Tadono et al., 2014), using GDAL tools in QGIS (GDAL Development Team, 2024; QGIS Development Team, 2024a, 2024b), embedded functions in Google Earth Engine, and the TAGEE package (Safanelli et al., 2020). Terrain attributes included slope, aspect, northness, eastness, convexity, topographic position index and topographic diversity (Theobald et al., 2015), landforms (Iwahashi et al., 2018; Iwahashi & Pike, 2007; Theobald et al., 2015), and horizontal-, vertical- and maximum curvature.

Parent material: Parent material was represented by the lithology field of the 1:250,000 Geological Map of Namibia (Geological Survey of Namibia, n.d.), generalised geological units (Atlas of Namibia Team, 2022), and mineralogical indices derived from Landsat and ASTER imagery. The Landsat-based mineral indices (carbonate, ferrous minerals, rock-outcrop normalisation) followed methods adapted from the USDA Soil Survey Handbook (Kienast-Brown et al., 2017) and earlier work by Bodily (2005), Jensen (2005), Nield et al. (2007), Boettinger et al. (2008) and Stum et al. (2010). The ASTER-based indices (kaolinite, quartz, calcite, carbonates and mafic minerals) were computed using the Ahrari (2024) algorithm, based on mineral detection methods described by Rajendran & Nasir (2019).

Age: In the absence of appropriate data, a direct soil age factor was omitted, though pedogenic maturity is implied by some of the other covariates, e.g. mineralogical indices and soil type.

Neighbourhood: Spatial hydrological predictors included distance to drainage lines and variables describing flow pathways and sediment redistribution, derived from MERIT Hydro (Yamazaki et al., 2019) and HydroSHEDS (Lehner et al., 2008; Lehner & Grill, 2013).

Terrain attributes and spectral indices were calculated within GEE. Where necessary, rasters were converted to floating-point format, reprojected to WGS84 (EPSG:4326), and resampled to 90 m spatial resolution. A 2 km buffer was applied during derivation of terrain attributes and vegetation indices to minimise edge effects, and the products clipped to the national extent after processing. Details and thumbnail maps of covariates are provided in Appendices B and C.

4.5.5 Covariate selection

Consistent with recommendations by [Kuhn & Johnson \(2013\)](#), [Brungard et al. \(2015\)](#), and [Kienast-Brown et al. \(2017\)](#), a hybrid feature-selection strategy was adopted, rather than relying solely on RF's internal variable ranking. Exploratory manual inclusion/exclusion screening was followed by testing of two automated methods. Recursive Feature Elimination (Guyon et al., 2002), implemented via the *caret* package in R (Kuhn, 2008, 2024) has been successfully applied in DSM (Kassai et al., 2025; Simperegui et al., 2025) and performed well in iteratively removing uninformative predictors. However, the *Boruta* algorithm (Kursa & Rudnicki, 2010) consistently yielded higher validation accuracy and was therefore chosen as the primary selection tool. Boruta operates as a wrapper around RF by generating shuffled “shadow” copies of each predictor and using a Z-score test to determine whether the real predictor provides more information than its shadow counterpart. Predictors failing this test are discarded iteratively (Kursa & Rudnicki, 2010). This study's process allows for manual override of the automated covariate selection.

4.5.6 Hyperparameter tuning, model training and validation

Predictor values from all environmental covariates were extracted at the georeferenced soil observation points and paired with the corresponding measured soil property values to form the regression matrices used for model training. Hyperparameter optimisation was undertaken to improve model accuracy and generalisability and to reduce overfitting. Model sensitivity to the parameters was assessed through initial manual exploratory tests, whereafter an automated tuning procedure was adopted to deal with the large number of models (16 soil properties x 3 depth intervals x up to 65 predictors).

Hyperparameters were optimised sequentially for each soil property–depth combination using the *tuneRanger* R package ([Probst et al., 2019b](#)), following approaches applied in recent DSM studies (Kassai et al., 2025; Simperegui et al., 2025). The parameters tuned were:

- *ntree*, the number of decision trees in the forest;
- *mtry*, the number of predictors randomly sampled at each split;
- *nodesize*, the minimum terminal node size to avoid overfitting;
- *sampsize*, the bootstrap sample size for each tree; and
- *max_Nodes*, the maximum number of terminal nodes per tree, which regulates model complexity.

The optimisation objective was to minimise Root Mean Square Error (RMSE) while maximising the Coefficient of Determination (R^2), finding the optimum model stability and predictive accuracy.

To train and validate the models, the dataset for each soil property and depth interval was repeatedly partitioned using 20 bootstrap iterations, with each iteration randomly splitting the data into 80% training and 20% validation sets. Bootstrapping was selected after testing k-fold and stratified cross-validation approaches, as it provided a stable estimate of predictive performance while accommodating the uneven spatial distribution of observations. All iterations were executed using a fixed random seed to ensure reproducibility. The results were 960 individual models (20 bootstraps $\times$ 16 soil properties $\times$ 3 depth intervals). For each iteration, statistical model performance was evaluated using RMSE to quantify prediction error and R^2 to assess explained variance. Summary statistics across the 20 iterations were computed to provide estimates of model performance and to enable comparison with national and international DSM studies. Lin's Concordance Correlation Coefficient (CCC) was calculated for the mean of the 20 iterations, to evaluate the agreement between observed and predicted values, accounting for both precision and bias (Lin, 1989).

4.5.7 Spatial predictions and uncertainty estimation

Each bootstrap model produced a spatial prediction across Namibia at 90 m resolution. The final spatial prediction for each soil property and depth interval was calculated as the per-pixel mean of the 20 bootstrap outputs, providing a defensible central estimate less sensitive to sampling variability.

Uncertainty estimation followed established digital soil mapping conventions (Arrouays et al., 2014; Heuvelink, 2018; Malone et al., 2017; Vaysse & Lagacherie, 2017). Within the GEE environment, the 5th and 95th percentiles of the bootstrap predictions were computed pixel-wise to represent the lower and upper uncertainty boundaries. The difference between these percentiles defines the 90% prediction interval (PI90), which captures both model error and resampling-induced variability. Maps of the mean, PI90 width, lower and upper limits were generated to visualise spatial patterns of uncertainty across Namibia.

To evaluate the calibration of these uncertainty estimates, the Prediction Interval Coverage Probability (PICP) was calculated (Malone et al., 2011; Shrestha & Solomatine, 2006). PICP represents the proportion of observed values that fall within the model-derived prediction interval. Well-calibrated uncertainty estimates should yield PICP values close to 90% for PI90, whereas substantial deviations indicate underestimation or overestimation of uncertainty (Poggio et al., 2021; Schmidinger & Heuvelink, 2023).

4.5.8 Pedological plausibility evaluation

Hengl et al. (2017b), Poggio et al. (2021), and Rossiter et al. (2022) recommended that statistical validation be supplemented by qualitative comparison of predicted spatial patterns with known soil–terrain relationships, climatic gradients, geological units, existing national and local soil maps, and personal knowledge from field experience to ensure pedological plausibility. Predicted soil property distributions and depth trends were assessed for consistency with dominant pedogenic processes in arid and semi-arid environments, including carbonate accumulation, low

leaching, aeolian sand deposition, textural differentiation, and organic matter dynamics. Special attention was given to spatial coherence within and across major physiographic units, as well as to abrupt discontinuities and unusual patterns, potentially introduced by covariate artefacts.

4.5.9 External validation

The Namibian analytical dataset generated under the Horizon 2020 *Soils4Africa* Project was excluded from model calibration and prediction. As a result, it provided an independent dataset for external validation and for benchmarking model performance against existing global and continental soil products. Of the Namibian samples, 63 analyses from 33 locations were obtained through conventional wet chemistry, while values for a further 131 samples were modelled from mid-infrared (MIR) spectral measurements (Gelsleichter & Coetzee, 2026).

Predicted soil property values corresponding to the *Soils4Africa* sample locations were extracted using the *terra* package in *R* (Hijmans, 2025) from the maps produced in this study, as well as from SoilGrids250m V2 (Poggio et al., 2021) and the iSDA soil products (Miller et al., 2021). Because soil properties were reported in different units across datasets, all values were standardised to the units used in this study prior to analysis. Differences in depth intervals between the present study, SoilGrids, and the *Soils4Africa* dataset required harmonisation. Predicted values were interpolated to the *Soils4Africa* reference depths (0–20 and 20–50 cm) using a mass-preserving equal-area spline (the *ea_spline* function in the *ithir* *R* package) (Malone, 2023), ensuring comparability while conserving the original depth-integrated quantities.

4.5.10 Mapping pipeline

The full mapping pipeline is presented in Appendix D. It proceeded through sequential stages, namely (1) data treatment — soil data cleaning and depth harmonisation; (2) feature selection and RF hyperparameter tuning in *R*; (2.5) covariate assembly and core RF modelling in GEE (including bootstrapping and uncertainty predictions); (3) post-processing of GEE outputs (including application of physically meaningful min/max limits, computation of PI90, and generation of cartographic plots); (4) computation of PICP90; (5) spectral modelling (optional; needed for the independent *Soils4Africa* dataset); and (f) independent/external validation (Gelsleichter, 2026b; Gelsleichter & Coetzee, 2026).

4.6 Integration — the Namibian Soil Information System (NamSIS)

The observational, spatial, and model-derived components were structured for integration into NamSIS through a shared geospatial reference system, standardised data formats, and relational linkages. The NSPD forms the core. These georeferenced point observations underpin both the national soil map, which provides the spatially explicit interpretation of soil-landscape relationships, and the modelling outputs, which include spatially continuous, qualitative predictions of soil properties. The components can be functionally interconnected through a user interface. Soil profile data inform both map unit delineation and model calibration. The national soil map serves as an environmental covariate and provides context for interpreting DSM outputs. In turn, the DSM products offer quantitative detail and reveal spatial patterns that can guide supplementary sampling and inform future refinement of the conventional map. Conceptually,

NamSIS thus operates as a multi-layered system in which data storage, interpretation, and modelling are not treated as separate entities, but as interdependent components of a unified soil information architecture. This enables cross-referencing between datasets, supports both qualitative and quantitative analyses, and provides a foundation for iterative system improvement as new data and methods become available.

The study focused on the scientific and structural components of this system and did not extend to the development of full web-based dissemination tools (e.g. APIs, web services, or interactive dashboards). Such functionalities represent a subsequent stage of system implementation. Nevertheless, the data structures, formats, and metadata standards established here are designed to support future integration into web-based platforms and broader spatial data infrastructures.

5. RESULTS AND DISCUSSION

The Namibian Soil Information System is the first systematic consolidation of Namibia's soil knowledge. Its components are elaborated in sections 5.1–5.4 and the integrated system presented in 5.5.

5.1 Legacy data preservation and documentation

One major objective of the study was the systematic recovery, preservation and organisation of Namibia's legacy soil information. Previously disconnected and at-risk materials were consolidated into a structured suite of archives, safeguarding irreplaceable historical information.

5.1.1 Original point (profile) data archive

The *ORIGINAL POINT DATA* archive contains the primary datasets generated during the 1998–2000 National Soil Survey, namely NSD1000, NSD250, NSDkava, and NSDowa. The NARIS Soil Dataset, representing the consolidated state of these and the MAWF–AEZ data by the year 2009, is preserved as an MS Excel file.

Additional source datasets are stored in their original formats, as spreadsheets, laboratory reports, field notes, and information embedded within project documentation. These include AMP, AMP–NA, AMP–NDLS, BIOTA, DAM, ETO, GIZ–Bush–SOC, KB, KH, LandPKS, LDN, LDN–Omusati, NH, SASSCAL, S4A, and TFO datasets. Where possible, links to external repositories are provided to maintain provenance and facilitate future updates.

The archive currently includes scans (PDF format) of the majority of NSS and MAWF–AEZ field forms. Field forms associated with the AO, DOR, MAR, NDONGA, OMAH_A, OMAH_B, ONGO, OVI, SPERR, and TSUM datasets could not be located and are presumed lost.

5.1.2 Analytical methods

Analytical procedures were compiled into an *ANALYTICAL METHODS* archive. Approximately 78% of legacy samples were analysed by the Ministry of Agriculture's AgricLab, 16.5% by the University of Hamburg, and the remainder by the Okavango Research Institute, the University of Pretoria, and the South African Institute for Soil, Climate and Water (Figure 17. Percentage of samples per laboratory in the Namibian Soil Profile Database (NSPD).Figure 17).

The archive includes:

- AgricLab standard operating procedures;
- GLOSOLAN Walkley–Black soil organic carbon (FAO/GLOSOLAN, 2019);
- AgriLASA (2004) laboratory methods;
- Soils4Africa project wet chemistry analyses guidelines (Paterson, 2021).

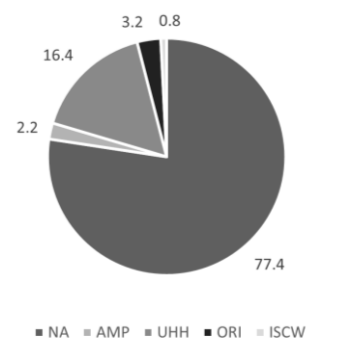


Figure 17. Percentage of samples per laboratory in the Namibian Soil Profile Database (NSPD).

5.1.3 Documents archive

The *DOCUMENTS* archive contains the bibliography and Zotero catalogue (https://www.zotero.org/groups/6649154/soil_literature_-_namibia/library) compiled during the study, covering published articles, dissertations, survey reports, and other grey literature related to Namibian soils. Where permitted, full-text documents are stored as PDF files; where copyright or subscription restrictions apply, references or external access links are provided.

To ensure long-term public access and spread the risk of information loss, digital documents are also archived outside NamSIS in two national repositories:

- Environmental Information Service (EIS) eLibrary (<http://the-eis.com/elibrary/>);
- Namibia University of Science and Technology (NUST) Library Special Collections (<http://specialcollections.nust.na:8080/greenstone3/library/collection/soils/page/about>).

5.1.4 Photograph archive

The *PHOTOS* archive contains more than 900 site and soil profile photographs linked to corresponding profiles within the NSPD and the GIS. Images are organised by dataset (Table 2 **Error! Reference source not found.**) and provide visual context for morphological interpretation, classification, and quality verification.

Table 2. Number and types of photos by dataset in the Namibian Soil Profile Database (NSPD).

Dataset	n	Type
BIOTA	157	Site, profile
HEIDE	7	Profile
NSS 1M	64	Profile
NSS OKA	43	Profile
NSS OWA	27	Profile
NSS TWO	11	Profile
OMAH-A	16	Profile
ONGO	13	Profile
S4A	340	Site, profile
SASSCAL	54	Site, profile
SPERR	24	Profile
TFO	148	Site, profile
TSUM	12	Profile

5.1.5 Original maps archive

The *ORIGINAL MAPS* archive comprises scanned and digitised legacy maps of soil, agro-ecological zones, land systems, and related thematic maps. These materials provided the cartographic basis for the development of the national soil map and are preserved as raster or vector files.

5.2 The Namibian Soil Profile Database (NSPD)

At the core of NamSIS lies the newly compiled Namibian Soil Profile Database (NSPD) of integrated, harmonised legacy soil observations, which forms the empirical foundation for national soil mapping and digital soil modelling. It has been designed to support data storage, retrieval, analysis, and future expansion.

5.2.1 Database structure

The NSPD relational database (Figure 18; Appendix F) is currently in MS Access format, with linked tables for profile registration, site descriptions, horizon field observations, horizon analytical data, depth-harmonised analytical data (3 depths), depth-harmonised analytical data (6 depths), environmental information, and metadata. The NSPD is also presented as an MS Excel workbook with four worksheets — *Reg&Site*, *Hor_lab*, *Hor_field*, and *Metadata* (see Appendix G) and as CSV files.

Each observation point is uniquely identified through the Profile Identification (PRID) field, which links site information to associated horizon descriptions and analytical data. Genetic horizons or standardised depth layers are identified by the Horizon Number (HONU) to link field descriptions to laboratory measurements. Datasets (projects or campaigns) are linked to metadata, photographs of sites and soil profiles, and publications or project documents.

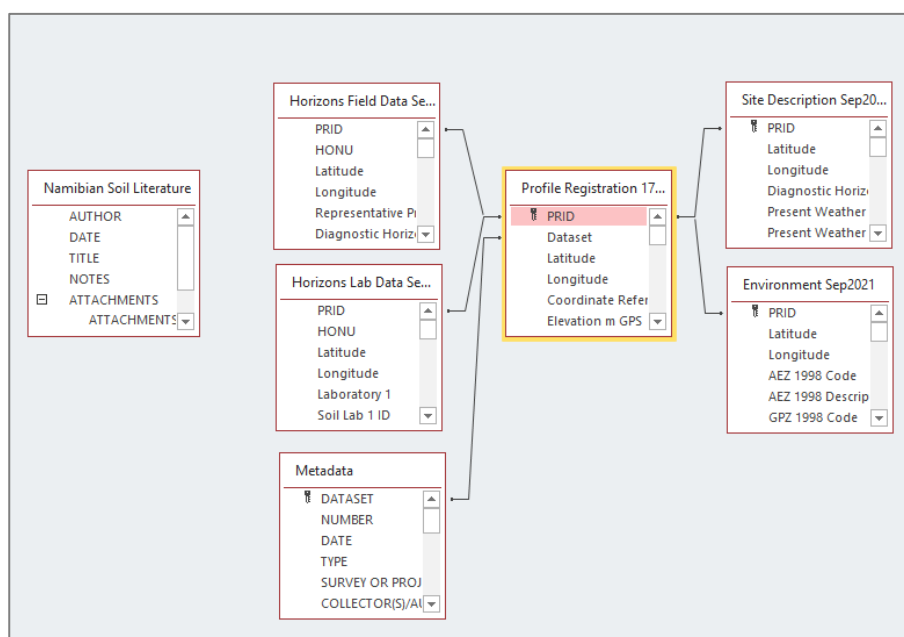


Figure 18. Relational structure of the Namibian Soil Profile Database (NSPD).

5.2.2 Data volume and dataset composition

NSPD currently contains 5 099 georeferenced observation locations (profiles, augerings, and sampling sites), associated with 13 425 horizons or depth layers with field and analytical data. This more than doubles the previously databased records from the 1998–2000 National Soil Survey — 2 155 points and 6 204 horizons (Figure 19) — in addition to a substantial improvement in spatial distribution. The expansion is the result of recovery of previously uncaptured information from the NSS field forms, as well as the integration of 30 other project datasets.

Through this data collation and standardisation effort, the Namibian holdings in the World Soil Information Service (WoSIS) increased more than twenty-fold, from 65 profiles — primarily originating from NamSOTER — to 1 730 profiles in 2026.

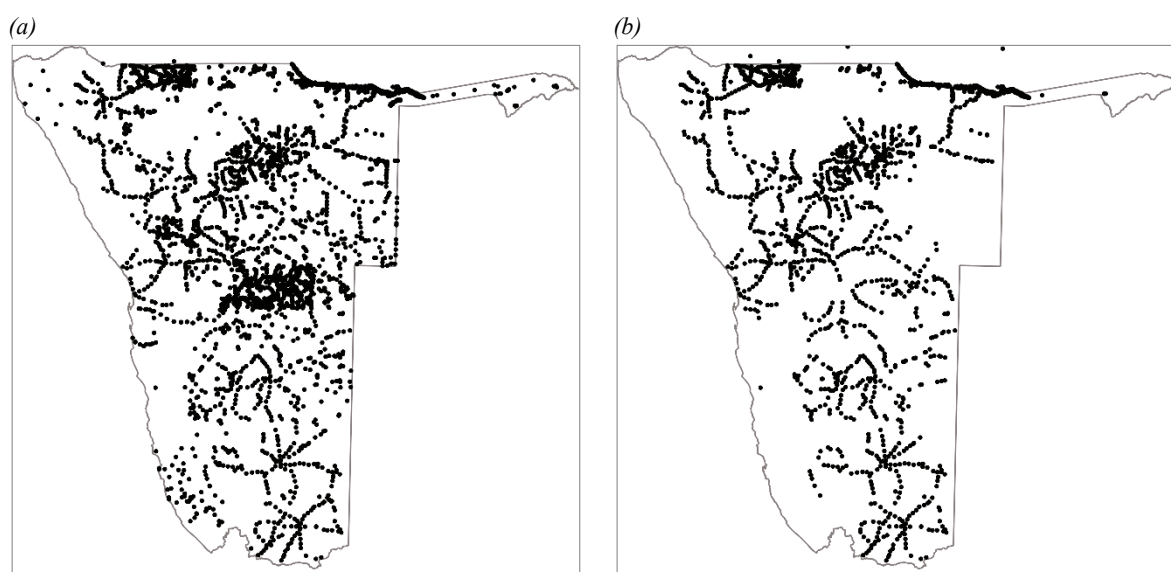


Figure 19. Distribution of soil point data in (a) the Namibian Soil Profile Database (NSPD) and (b) the 1998-2000 National Soil Survey of Namibia.

The constituent datasets and number of observation locations within each are summarised in Table 3, while spatial distribution maps per dataset are presented in Appendix B.

Table 3. Current datasets in the Namibian Soil Profile Database (NSPD) and their respective number of sites (profiles, augerings, or sampling points). Abbreviations/acronyms are clarified in paragraph 4.3.3.1 and the NSPD metadata.

Dataset	n	Dataset	n	Dataset	n
AMP	62	KH	14	ONGO	21
AMP-NA	108	LandPKS	918	OVI	37
AMP-NDLS	9	LDN	222	OWA NSS	309
AO	69	LDN-Omusati	106	SASSCAL	30
BIOTA	257	MAR	19	Soils4Africa	112
DAM	28	MAWF-AEZ MISC	61	SPERR	44
DOR	59	NAM NSS	828	TFO	141
ETO	54	NDONGA	17	TSUM	16
FEB	86	NH	48	TWO NSS	417
GIZ-Bush-SOC	162	OKA NSS	621	Uitkomst	4
JAN	52	OMAH_A	27		
KB	35	OMAH_B	79		

The level of detail within the NSPD varies considerably between records. Some observations contain only basic site information and field observations, whereas others include full profile descriptions with extensive laboratory analyses. This heterogeneity is the result of differences in original survey objectives and budget constraints.

Table 4 summarises the number of laboratory records available for key physical and chemical soil properties, illustrating the uneven distribution of analytical data across variables.

Table 4. The current number of laboratory records per soil property in the Namibian Soil Profile Database (NSPD).

Property	n	Property	n	Property	n
Sand	4 702	Mg exch + sol (Mehlich)	125	Cu	63
Coarse sand	63	Mg sol (NH ₄ acetate)	1 458	S	63
Medium sand	63	Mg exch	586	Zn	63
Fine sand	63	K sol (NH ₄ acetate)	1 457	Fe_ox (oxalate extract)	63
Very fine sand	63	K exch + sol (Mehlich)	125	Al_ox (oxalate extract)	63
Silt	4 064	K exch	585	inorgC	63
Coarse silt	63	Na exch + sol (Mehlich)	125	totC	63
Fine silt	63	Na sol (NH ₄ acetate)	1 449	orgC	63
Clay	4 064	Na exch	587	exchH	63
Silt + clay	638	OC (Walkley-Black)	4 984	exchAl	63
Textural class	4 073	OC (LOI)	162	exchAcid	63
Tex (field)	3 353	P total (Mehlich)	180	totMo	63
pH water	5 000	P (Olsen)	1 823	totCd	63
pH KCl	103	N total	1 504	totPb	63
pH CaCl ₂	1 469	Total CO ₃ eq	430	totV	63
Bulk density	1 420	Sol CO ₃	50	totHg	63
EC 1:2.5	4 032	Gypsum	79	totCr	63
EC sat paste	1 748	Sulfate	214	totCo	63
CEC eff calc	813	Chloride extractable	162	totNi	63
CEC measured	928	Fe (Mehlich)	180	totCu	63
Base sat est (pH)	3 037	Al (Mehlich)	180	totZn	63
Base sat est (CEC)	517	Mn (Mehlich)	180	totAs	63
Ca exch + sol (Mehlich)	125	Porosity	54	totSb	63
Ca sol (NH ₄ acetate)	1 460	CaCO ₃	63		
Ca exch	585	B	63		

5.2.3 Pedogenic horizons, fixed depth layers, and depth-harmonised data

The horizon tables of the NSPD contain pedogenic horizons or predefined depth layers. For example, LandPKS observations were recorded at 0–1, 1–10, 10–20, 20–50, 50–70, 70–100, and 100–150 cm. These layers were assigned consecutive horizon numbers. Master (e.g. A, E, B, C) and diagnostic horizons were recorded separately.

Two separate tables contain the depth-harmonised analytical data: one for three agriculturally relevant depths (0–30, 30–60, 60–100 cm), and one for the six standard GlobalSoilMap / SoilGrids depths (0–5, 5–15, 15–30, 30–60, 60–100, 100–200 cm). After the addition of new analytical data, these tables can be updated by re-running the R scripts.

5.3 The National Soil Map of Namibia (NSMN)

5.3.1 Overview

Conventional soil mapping constitutes the expert-driven component of NamSIS, translating harmonised legacy observations and pedological interpretation into categorical soil map units. The principal outputs from the conventional soil mapping process are:

- The National Soil Map of Namibia (NSMN) at a nominal scale of 1:250 000, with soil map units comprised of dominant, associated and included soil typological units (Figure 20 and

Figure 21);

- Spatial distribution maps of individual WRB Reference Soil Groups, derived from classification qualifiers (Figure 24);
- Spatial distribution maps of key diagnostic soil attributes, derived from classification qualifiers (Figure 25); and
- Maps of terrain attributes and land systems, generated during the production of the CSM (Figure 26 to Figure 32).

All maps are available as digital files in shape or GeoTIFF format and are thus updatable in a geographical information system (GIS). Future revisions can selectively increase spatial detail in newly sampled regions without compromising overall consistency.

Figure 20 shows the distribution of the 13 dominant WRB Reference Soil Groups (RSGs) at the highest classification level: Arenosols, Leptosols, Regosols, Calcisols, Cambisols, Gypsisols, Solonchaks, Solonetz, Fluvisols, Durisols, Gleysols, Vertisols, and Technosols. Full spatial and typological details can be accessed within the GIS, including local pockets of Luvisols, Stagnosols, Planosols, Histosols, and palaeo-Kastanozems.

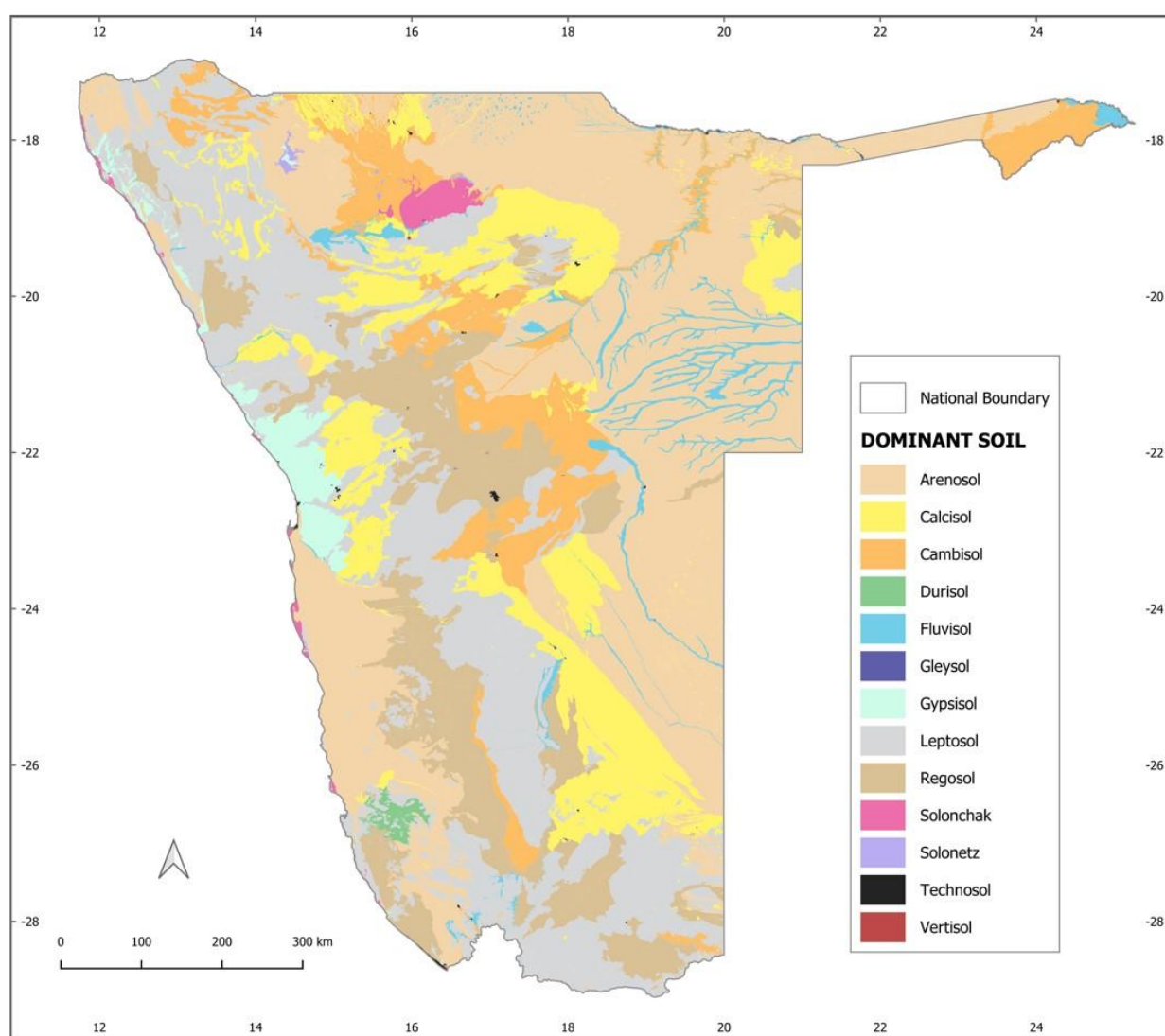


Figure 20. The new National Soil Map of Namibia — dominant World Reference Base (WRB) Reference Soil Groups (RSGs).

The NSMN map consists of 4787 polygons (Figure 21) forming 490 soil map units (SMUs) composed of 212 soil typological units (STUs) (Appendix H). An SMU represents a combination of one dominant STU (covering at least 50% of the polygon area), an associated STU (covering 25–50%) (if applicable), and up to two included STUs (covering 5–25%). The minimum mappable unit is in the order of 1 km², although such small units were generally only delineated where soils or landforms were very distinctive, for example pans and inselbergs. This taxonomic structure allows users to work both with dominant patterns and with important subdominant or minor soil components that may be locally significant for land use. The full legend, including the complete WRB classification of each map unit, is provided in the appendices and digital GIS data.



Figure 21. The new National Soil Map of Namibia (NSMN) — outlines of all soil map units.

The new coverages of the three areas that were left unmapped in the Preliminary Soil Map (Coetzee, 2001), are displayed in Figure 22.

Compared to the NSS 1M map ((ICC & MAWRD, 2000), the new NSMN captures substantially greater spatial heterogeneity (Figure 23).

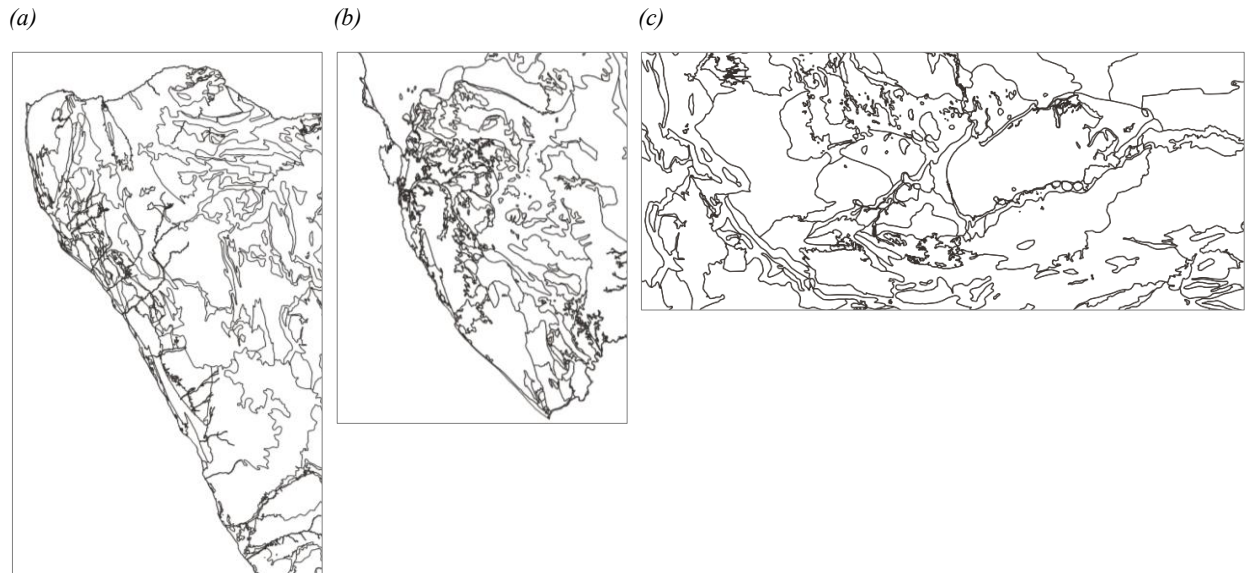


Figure 22. Excerpts from the new National Soil Map of Namibia showing the coverages of previously unmapped areas unmapped in the NSS 1M map (ICC & MAWRD, 2000): (a) Skeleton Coast National Park, (b) Tsau ǁKhaeb National Park, and (c) Etosha National Park.

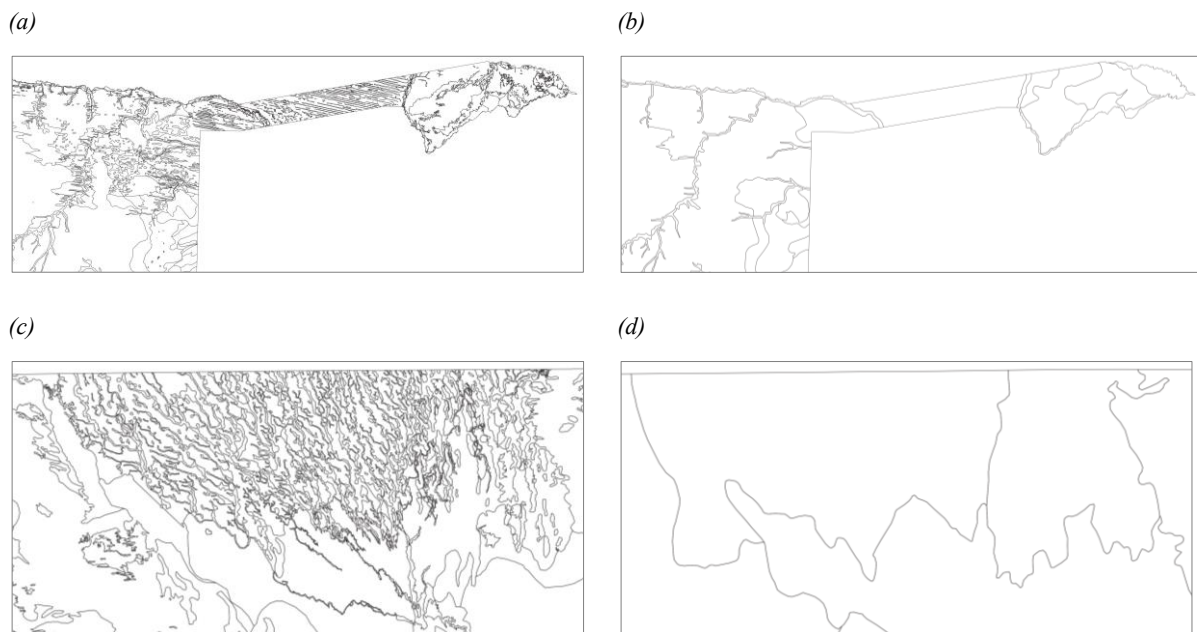


Figure 23. Excerpts from the new National Soil map of Namibia ((a) and (c)) and the NSS 1M Map ((b) and (d)) to illustrate the increase in spatial detail: (a) and (b) northeastern Namibia; (c) and (d) north-central Namibia.

5.3.2 Distribution of RSGs, environmental controls and implications for land use and management

Beyond its technical and cartographic advances, the NSMN enables clearer interpretation of soil–landscape relationships and distribution of soil constraints and potentials in a manner that is directly relevant to land management at national and regional scales.

Large portions of Namibia are dominated by shallow and stony *Leptosols*. These landscapes impose limitations on mechanised cropping due to restricted rooting depth, low plant-available water storage, and physical barriers to tillage. *Leptosols* support shallow-rooted grasses and forbs, while shrub and tree roots will find joints in fractured bedrock. Their greatest potential lies in extensive land uses such as livestock farming, wildlife-based conservation, and tourism, where productivity depends more on the resilience of indigenous vegetation, grazing management, and access to drinking water than on agronomic potential.

Namibian *Regosols* are usually poorly developed and highly erodible, with low water-holding capacity and physical coherence. As a result, their agronomic potential is low, but they provide valuable grazing for wildlife and livestock.

Dominating the crop producing areas of Namibia, the low water- and nutrient retention of *Arenosols* have impacts on both dryland and irrigated agriculture. With careful management of vegetation cover, organic matter, soil structure, and rainwater harvesting, *Arenosols* can support rainfed production of grains, tuber and root crops, melons, legumes, and fodder crops where rainfall exceeds 350–400 mm per year. Where irrigation — or even just supplementary irrigation — is possible, *Arenosols* can be quite productive, using frequent, efficient water application through drip and sprinkler systems, and staged nutrient inputs to reduce leaching losses. Fine sands are prone to compaction by machinery, livestock trampling or inappropriate tillage. Fine to medium-grained sands are highly susceptible to wind erosion if vegetation cover is lost, and overgrazing can trigger dune remobilisation. In some *Arenosols*, hydrophobic behaviour due to fungal exudates impedes water infiltration, especially after prolonged hot, dry periods. The Namib Sand Sea, Skeleton Coast and Kunene dune fields are dominated by protic *Arenosols*. Interactions between fluvial and aeolian processes along the lower reaches of the Namib's ephemeral rivers determine where *Arenosols* develop and persist. Where fluvial processes dominate, such as along the Hoarusib, Kuiseb and Kunene Rivers, episodic floods flush sand from riverbeds and interrupt the northward migration of dune fields.

Cambisols are some of the comparatively better Namibian soils for cultivation. The weathering of parent material and accumulation of clay minerals enhance cation exchange capacity and nutrient retention, while soil structure — though weakly developed — increases porosity, internal drainage and water holding capacity.

Gypsisols in the Namib demonstrate how coastal atmospheric processes can leave a strong pedological imprint even where rainfall is minimal. Their distribution reflects the combined influence of hyper-aridity (where evaporation greatly exceeds precipitation), coastal fog influence, marine-derived sulfate inputs linked to upwelling in the Benguela Current, and the long-term geomorphic stability of coastal planation surfaces. Namibian *Gypsisols* are most strongly developed in the first 5–10 km from the coast, downwind of the strongest upwelling cells, and where fog incidence is highest. They are particularly well developed in the central Namib between the Kuiseb and Ugab Rivers, but less common in the northern and southern Namib where stronger wind regimes promote aeolian remobilisation of surficial materials, and they are absent from mobile dune fields. *Gypsisols* can be found as much as 50 km inland. The observed transition from gypsum-dominated horizons near the coast to calcic and petrocalcic horizons further inland reflects changes in moisture regime and solute supply. In the transition zone, calcic and petrocalcic horizons appear above gypsic and petrogypsic horizons. This is an indication that downward leaching is the main mechanism of accumulation in the Namib, as calcium sulfate is more soluble than calcium carbonate. The situation is reversed in the case of coastal saltpans or depressions

with a shallow water table: upward accumulation, driven by evaporation, results in the formation of the gypsic horizon above the calcic horizon. As Namibian Gypsisols are mainly found in the Namib Desert, agronomic potential is a moot point. They play a key role in the complex ecology of biocrusts and the stabilisation of gravel plains and are fundamental to the survival of endemic gypsophilic fauna and flora.

Durisols along the eastern margin of the central and southern Namib — particularly well expressed in the vicinity of Aus — provide another example of soils linked to old, stable surfaces and long-lasting geochemical conditions. Their association with alkaline environments and geomorphological stability supports the interpretation that duric and petroduric horizons are cumulative features requiring long timespans to develop. Lateral grading between duric, gypsic and calcic horizons observed in some areas highlights continuity among major pedogenic pathways in arid landscapes, rather than sharp categorical separations. *Durisols* can support grazing and drought-tolerant crops under irrigation where gypsum and calcium carbonate contents are not excessive in the upper horizons, provided that restrictive horizons are broken up (e.g. by deep ripping), and salinity and boron toxicity are managed. In Namibia, their location within arid to hyper-arid regions means that they mainly support wildlife and small stock under extensive ranching conditions.

Calcisols and calcrete formations are widespread throughout the country. Some are palaeosols, whereas others are actively developing in calcium- and magnesium rich alluvial, colluvial or aeolian materials. Namibian *Calcisols* allow crop production where rainfall is sufficient or irrigation water available, provided fertility, salinity, and erosion are managed. Many *Calcisols* have fine to medium-textured subsoils with relatively good water retention. Surface horizons are prone to slaking and crusting, which restricts infiltration and can increase runoff and erosion risk. Root penetration and internal drainage are impeded when continuous, strongly indurated petrocalcic layers occur close to the surface. However, a petrocalcic horizon beneath a thick sandy B horizon can be an asset in an arid climate, as it allows water to be retained in the root zone for longer. High calcium levels can restrict the availability of other cations to plants, and phosphorus and trace element uptake is often limited under such conditions. Some *Calcisols* are saline, and boron toxicity can occur where pH is high and boron is mobilised. Natural vegetation on *Calcisols* typically consists of nutritious shrubs, trees, ephemeral grasses and forbs, which provide good grazing, and soil faunal activity is usually high.

Solonchaks and *Solonetzs* are concentrated in endorheic basins like the Cuvelai-Etoshia system, coastal salt pans and terminal playas of ephemeral rivers. Their distributions reflect hydrological confinement, shallow groundwater tables, and long-term accumulation of salts driven by evaporation exceeding precipitation. Their frequent association with calcareous and gypsiferous horizons underlines that salinity, sodicity, carbonate and gypsum accumulation are often coupled in Namibian environments, and that soil constraints commonly occur in combinations. Many Namibian soils have raised salinity but do not meet the full WRB criteria for classification as *Solonchaks*. In such cases, salic and endosalic qualifiers are applied to other soil groups. *Solonchaks* and *Solonetzs* present severe impediments to crop production, especially in low-lying basins and confined drainage systems. Plants must overcome both matrix and osmotic potentials to extract water, and specific ions can be toxic or cause nutrient imbalances. Most desert, arid and semi-arid zone vegetation has some degree of salinity tolerance. For example, halophytic plants such as *Salicornia*, *Tamarix* and *Salsola* tolerate high soil salinity while *Mopane* trees can endure moderate salinity. Faunal activity is restricted in *Solonchaks* and mostly absent where the salt concentration exceeds 3%. *Solonetzs* pose additional physical challenges to water infiltration. Both

Solonchaks and Solonetz require specialised management, thus alternative low-intensity land uses are generally more realistic than costly reclamation measures.

Genetically young *Fluvisols*, intermingled with *Gleysols*, *Cambisols*, and *Arenosols*, are found in the floodplains of perennial rivers (e.g. the Okavango, Kwando–Linyanti–Chobe, and Zambezi), the Cuvelai-Etosa system, and in palaeodeltas of rivers discharging into the Atlantic Ocean. Anastomosing river systems like the Cuvelai drainage are characterised by very low gradients and low transport energy. As a result, they preferentially redistribute fine sediments, which accounts for the greater prevalence of silty, loamy, and clayey soils in their lower reaches and floodplains. These Fluvisols are generally fertile and relatively easy to work, with good potential for agriculture if not affected by shallow groundwater or salinity.

The riverbeds of ephemeral rivers in the interior are mainly composed of *Leptosols*, *Fluvisols* and *Calcisols*. Short-duration floods replenish their alluvial aquifers, providing water to deep-rooted vegetation, wildlife and humans in the dry season.

Luvisols are considered among the more productive soils in Namibia due to their relatively higher colloid content, favourable structure, high cation exchange capacity, good aeration, and moderate to high available water content. However, their spatial extent is limited to local inclusions, particularly in the north-central uplands and broad, shallow, usually dry river valleys (locally known as “omiramba”) in eastern Namibia.

Gleysols occur in waterlogged environments such as floodplains, marshes and seasonal wetlands, including the Cuvelai system and Lake Liambezi. They are also found in tidal zones and seasonal clay pans. Gleysols are saturated near-permanently, leading to anaerobic, reducing conditions which complicate their management. Paddy rice trials at Kalimbeza in Zambezi Region showed that their cultivation is agronomically possible, but unlikely to be profitable, and they are best left under natural wetland vegetation for use by wildlife and livestock.

Small inclusions of *Planosols* occur north of Etosha and in the Zambezi floodplains as seasonally waterlogged soils with abrupt textural contrasts.

Histosols are scarce in Namibia, although small patches of Fibric Histosols (Fluvic) occur along permanent rivers in Kavango-East and Zambezi Regions. Elsewhere, histic properties occur within other soil types without meeting the classification criteria for fully developed Histosols.

Vertisols are spatially restricted. Their occurrence in paleo-lacustrine basins and in fine-textured parent materials indicates that shrink–swell dynamics can develop where clay-rich sediments accumulate and seasonal wetting and drying cycles occur, even within an otherwise arid context. Even where full WRB Vertisol criteria are not met, the presence of vertic properties remains important for land management (for example, in the Maize Triangle between Otavi, Grootfontein and Tsumeb) and engineering considerations (such as dealing with cracking houses in Mariental, the site of an ancient lakebed).

Given Namibia’s relatively low urbanisation and industrial footprint, the spatial extent of *Technosols* is quite limited. The main types identified include Ekranic Technosols (sealed by technic hard materials such as asphalt and concrete), Urbic Technosols (containing anthropogenic rubble and refuse), Spolic Technosols (associated with industrial waste such as mine spoil, dredgings, slag, ash and rubble), and Garbic Technosols (containing organic waste).

5.3.3 Derived thematic maps

5.3.3.1 Categorical maps of individual soil classes

Thematic maps were generated to explore the spatial distribution of individual Reference Soil Groups in more detail (Figure 24). Each map indicates whether a given RSG is dominant ($\geq 50\%$), associated (25–50%), included (5–25%), or absent from each SMU.

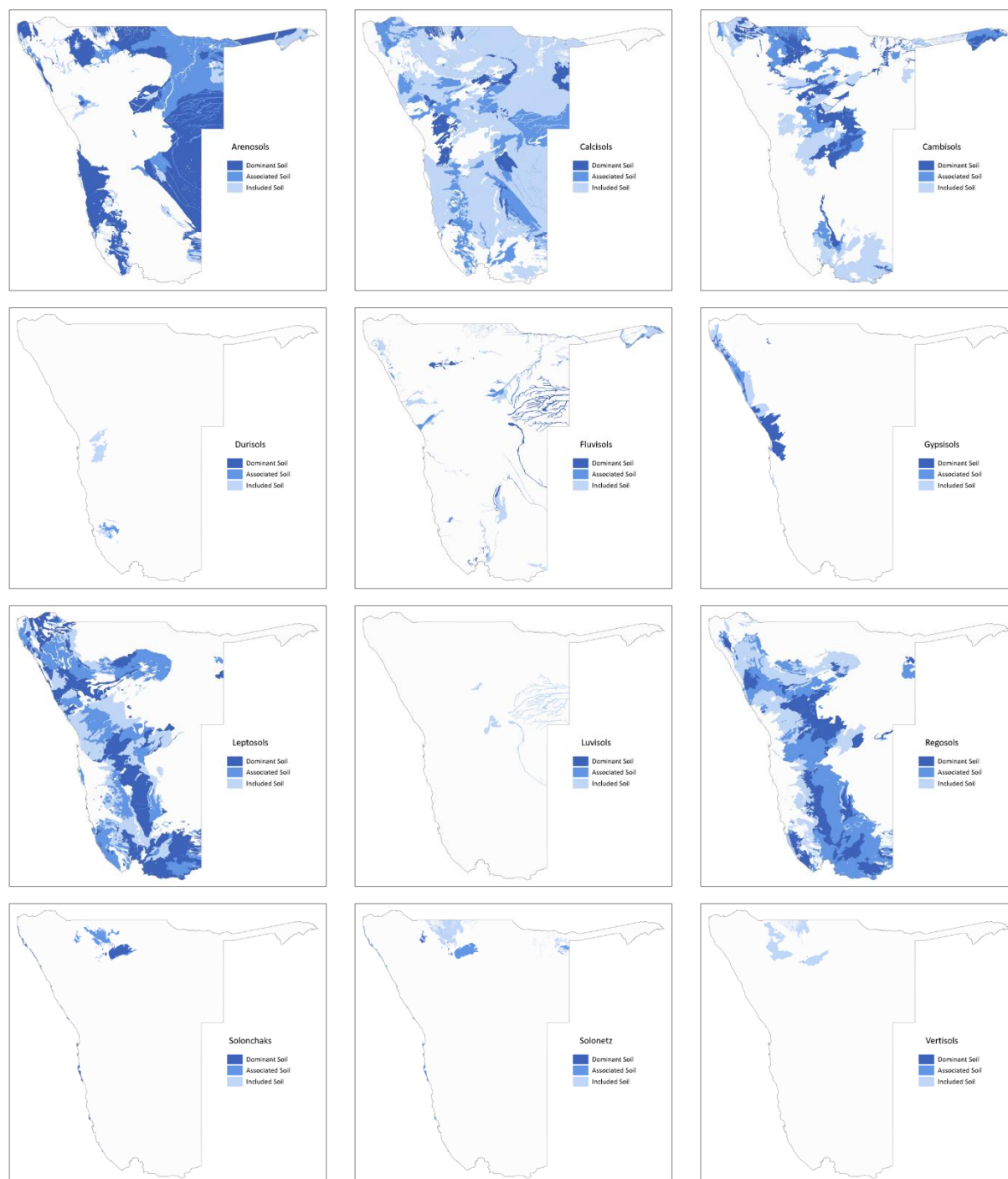


Figure 24. Distribution of major WRB RSGs to indicate where each one is spatially dominant ($\geq 50\%$) (dark blue), associated (25–50%) (medium blue) or included (5–25%) (light blue). (Full resolution maps can be viewed and downloaded at <https://soilsofnamibia.com/maps.html>)

These maps allow users to identify core geographic domains of each soil group, as well as transitional and mixed areas.

5.3.3.2 Categorical maps of selected soil properties

A set of thematic maps was generated to show the spatial distribution of key soil properties and diagnostic features, especially those defined by specific WRB qualifiers (Figure 25).



Figure 25. Distribution of selected soil properties and WRB diagnostic features, to indicate where each is spatially dominant ($\geq 50\%$) (dark green), associated (25-50%) (medium green) or included (5-25%) (light green). (Full resolution maps can be viewed and downloaded at <https://soilsofnamibia.com/properties.html>)

These maps complement the RSG maps by highlighting functional attributes such as shallow

depth, stoniness, calcrete development, dolomitic substrates, hydromorphic conditions and depositional environments.

Shallow soils — either Leptosols or having Leptic, Lithic, Nudilithic, or Lapiadic qualifiers — collectively delineate areas where continuous rock or technic hard material occur within 100 cm (Leptic), within 10 cm (Lithic), or at the surface (Nudilithic). Lapiadic soils — best developed in the Karstveld — combine Nudilithic conditions with distinctive dissolution features (“karren”) on limestone or dolomite surfaces. The shallow soils map reveals that bedrock lies near the surface across much of the Central Plateaus, Escarpment, Kaokoveld, Damaraland, the Karstveld and the far south, and around inselbergs and rocky ridges in the Namib Desert — covering about 46% of the country.

Stony soils were mapped using the Skeletic and Hyperskeletic qualifiers, indicating soils with more than 40% and more than 80% (by volume) coarse fragments, respectively, in the upper 100 cm or down to continuous rock or indurated horizons, whichever is shallower. These conditions are widespread in Leptosols and many Regosols of the Central Plateaus, Escarpment, Damaraland, parts of Kaokoland, and southern Namibia. The stoniness map highlights areas where shallow, coarse-textured soils inhibit tillage, root penetration and water storage, but often favour extensive grazing and wildlife habitat.

Dolomitic soils occur where parent materials are rich in dolomite ($\text{CaMg}(\text{CO}_3)_2$). They are prevalent in the Otavi Highlands, Karstveld, and along the southern margins of Etosha.

Sideralic soils exhibit low cation exchange capacity (below $24 \text{ cmol}(+) \text{ kg}^{-1}$ soil), or both very low CEC (below $4 \text{ cmol}(+) \text{ kg}^{-1}$ soil) and intense colour (a moist Munsell colour chroma of at least 5) in a subsurface horizon, reflecting enrichment in iron and/or aluminium oxides/hydroxides. Most Kalahari sands in Namibia are sideralic.

Hypercalcic soils have calcic horizons with more than 50% CaCO_3 equivalent, occurring as continuous but soft concentrations, in the fine earth fraction. *Petrocalcic* soils have a petrocalcic horizon within 100 cm of the surface — relatively continuously cemented or indurated nodular, massive, lamellar or petrified lamellar accumulation of secondary carbonates. These calcrete formations are widespread across Namibia.

Protic soils lack any signs of horizon development and are typical of active dune sands and other very young deposits, for example in the Namib Sand Sea.

Yermic soils are usually associated with desert pavements — gravel or stone fragments embedded in a loamy vesicular layer — that is frequently found in the upper few centimetres of unvegetated, structurally unstable silty loam soils in arid regions (Ellis, 1988, 1990). A common feature is polygonal networks of desiccation cracks that extend into the underlying layers and are filled with in-blown material. Desert varnish, ventifacts, soluble mineral accumulations, thin aeolian sand or silt may occur on the surface. Yermic soils occur on the desert gravel plains, and parts of Damaraland, the Kaokoveld, and in southern Namibia.

Colluvic soils are found where downslope movement under the influence of gravity results in the accumulation of poorly sorted heterogeneous material, usually of lower bulk density than the underlying buried soil. These typically occur on concave slopes, footslopes, depressions and fans, and are common in the Escarpment and Central Highlands of Namibia.

Fluvic soils develop in active or past riverine, lacustrine and coastal deposits, displaying distinctive, irregular stratification in texture, coarse fragments, colour and/or organic matter.

Stagnic soils are defined by seasonal surface water stagnation, and vertical changes in redox conditions. Distinctive reductimorphic mottling is present above a denser, slowly permeable subsoil. These conditions are rare in Namibia and mainly occur in small areas north of Etosha, and in the eastern Zambezi Region.

Gleyic soils indicate permanent or near-permanent groundwater saturation from a shallow or rising water table, often caused by lateral inflow from rivers. They typically occur in floodplains, valley bottoms, riparian zones and depressions with high water tables. Anaerobic conditions prevail for much of the year, resulting in gleyic properties (bluish-grey, greenish or mottled colours) and redoximorphic features caused by reduction of iron. In Namibia, small pockets occur in association with Fluvisols in the Cuvelai system, Okavango and Zambezi floodplains.

By mapping qualifiers, the NSMN framework provides a practical bridge between classification and function, translating WRB diagnostic criteria into spatial representations of soil constraints and potentials. This supports screening-level land evaluation and stratification, and helps prioritise where more detailed assessments, sampling, or project-specific mapping should be undertaken.

5.3.3.3 Terrain attributes

A suite of terrain attribute provides a physiographic framework for environmental interpretation and land-use implications.

Elevation (Figure 1) influences temperature regimes, frost risk and rainfall patterns, including orographic enhancement in the Otavi Mountains and Maize Triangle, and rain-shadow effects on the leeward side. Higher elevations, such as the Khomas Hochland, experience lower temperatures and higher solar radiation (Schulze & Horan, 2007). Cold air drainage into valleys increases frost incidence under certain winter conditions. Research has shown that even moderate elevations (1 200–1 600 m) can affect livestock health, productivity and reproduction (Pauling et al., 2018; Titaouine & Meziane, 2015).

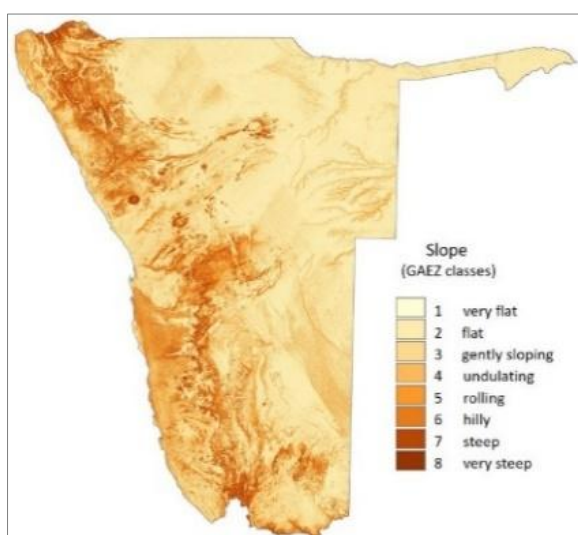


Figure 26. Slope Map of Namibia (GAEZ classes).

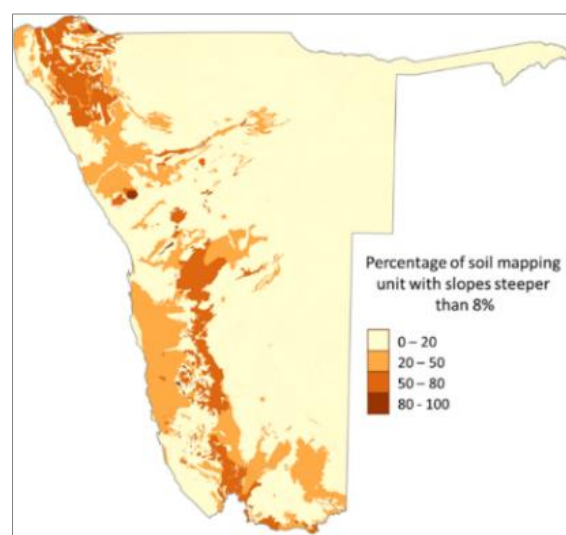


Figure 27. Percentage of each SMU with slopes steeper than 8% in Namibia.

Slope gradient is highest along the Escarpment and in the central highlands (Figure 26), where both sheet and gully erosion are active. In contrast, the north-central, northeastern and eastern regions are predominantly flat to gently sloping, posing few topographic barriers to mechanised

agriculture and exhibiting low water erosion risk (Figure 27). *Aspect* is less critical for Namibian agriculture because the main crop-producing areas are both in the tropics and mostly flat to level. However, aspect plays an important ecological role through microhabitat differentiation within the coastal fog belt, which is influenced by prevailing south-west winds. This is evident in the proliferation of lichens and specialised vegetation on western and south-western faces of ridges and inselbergs.

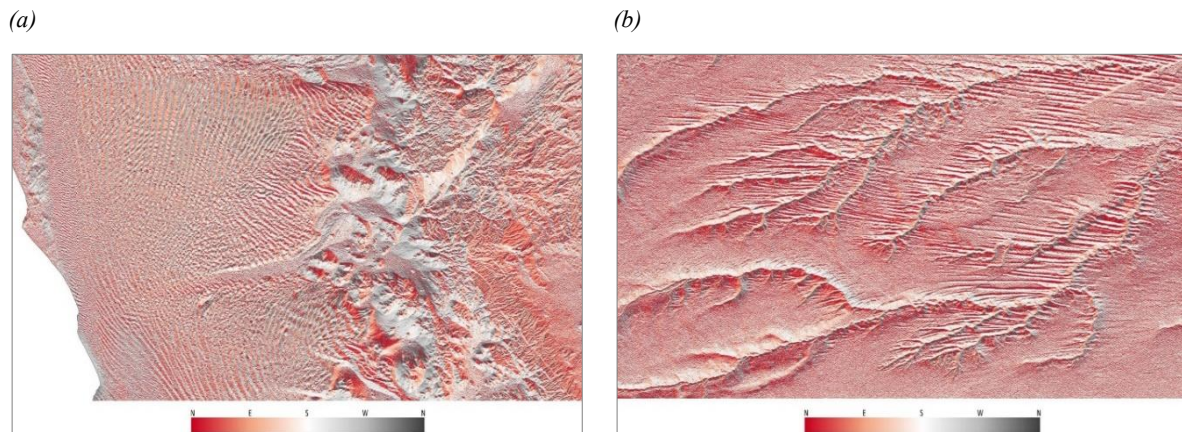


Figure 28. Aspect maps of Namibia: (a) Dune orientation highlighted in the Sossusvlei area and part of the Namib Sand Sea; (b) Subtle omiramba drainage patterns in the Epukiro-Eiseb area.

Aspect maps (Figure 28) and combined *aspect-slope* maps (Figure 29) proved valuable for soil mapping by highlighting dune orientations and aiding the detection of inselbergs, gorges, scarps, subtle drainage patterns, and other low-gradient landscape features such as lunette dunes along pan margins.

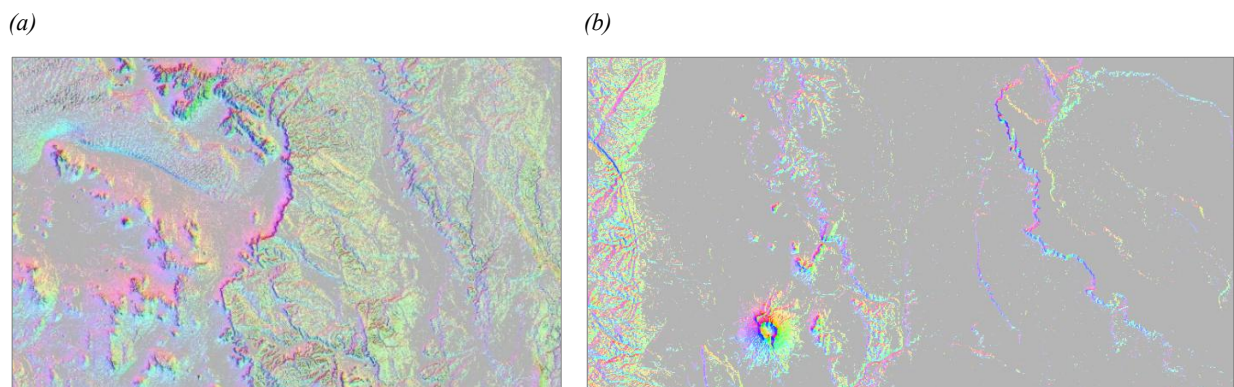


Figure 29. Combined aspect-slope maps of Namibia: (a) Inselbergs in the Namib, river gorges, the Schwarzrand scarp, and the abrupt transition from the rugged escarpment to the coastal plain and Namib Sand Sea; (b) Brukkaros Mountain, the Weissrand scarp, river valleys cutting through the Kalahari Sands, and the transition from the relatively level interior plateau to the incised escarpment.

Incised landscapes, highlighted by the *Riley Terrain Ruggedness Index (TRI)* (Figure 30), occur in the moderately to highly rugged terrain of the Khomas Hochland, northern Kunene Region, Escarpment, and parts of southern Namibia. This type of terrain affects animal husbandry, as it provides concealment for predators, increase livestock energy expenditure, and complicate herd management.

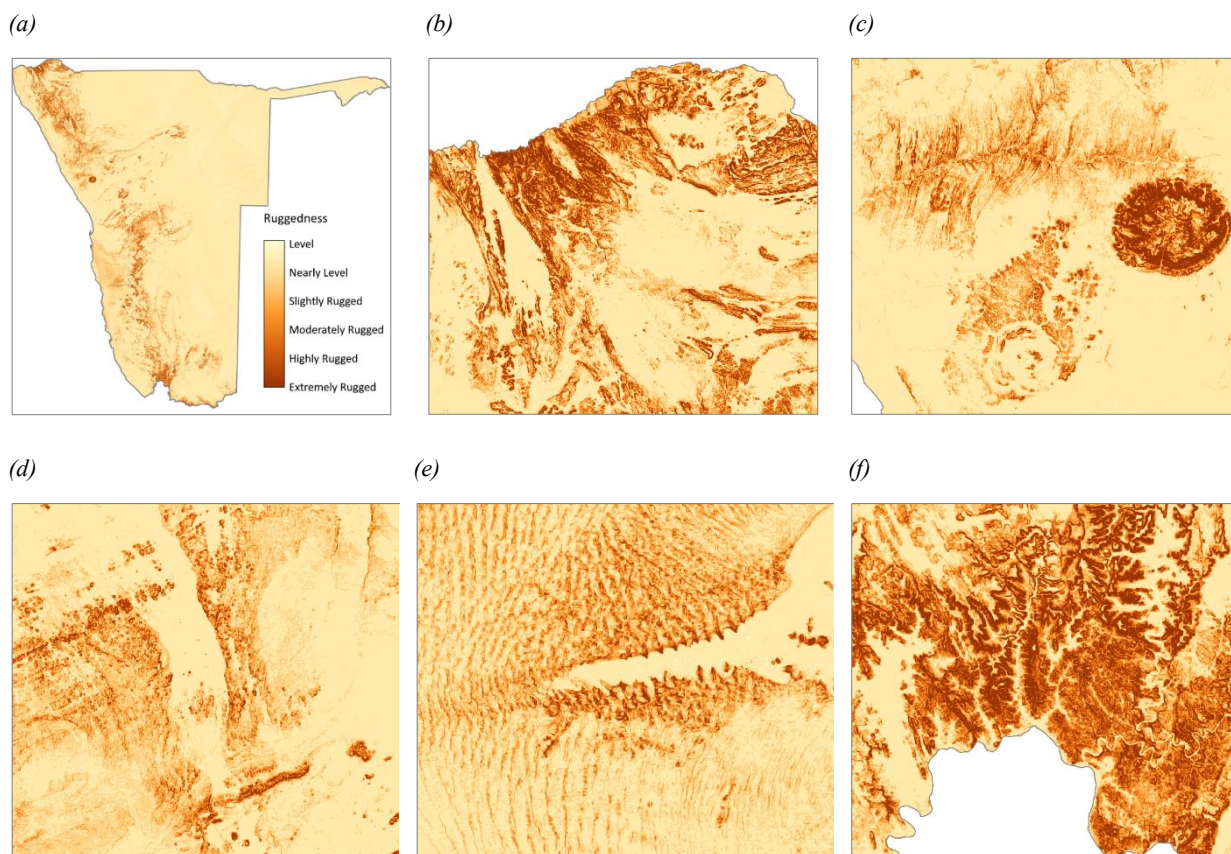
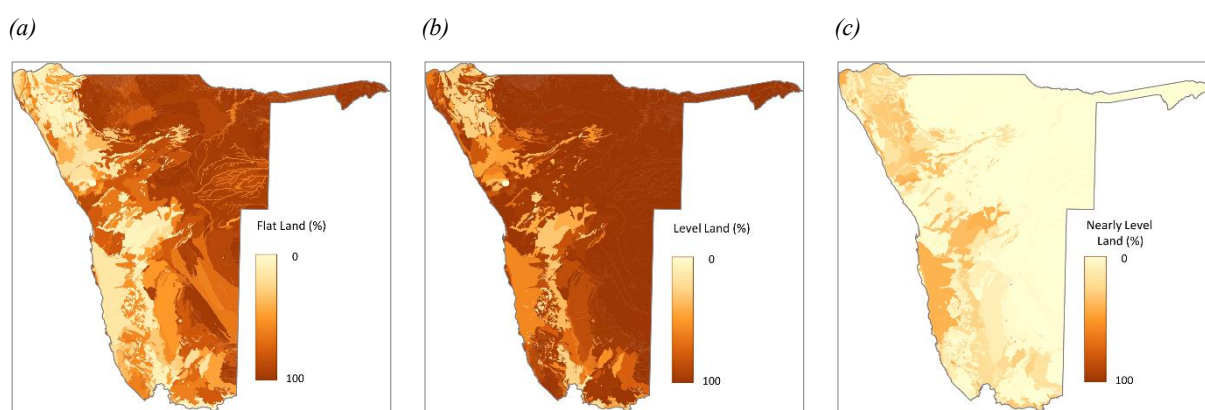


Figure 30. Terrain Ruggedness Index (TRI): (a) Namibia; (b) Marienfluss–Epupa area; (c) Brandberg–Messum–Ugab River area; (d) Windhoek Valley, Khomas Hochland, and Auas Mountains; (e) Sossusvlei area, highlighting star and transverse dunes, the valley of the Tsauchab River and terminal playas; (f) Aussenkehr area and the confluence of the Fish and Orange Rivers.

Highly dissected terrain reduces the suitability of land for mechanised agriculture. However, this is not a significant hindrance in Namibia, as analysis of TRI values by soil map unit (Figure 31) shows that flat and level terrain predominates in the northeast — the region with climatic conditions most suitable for rainfed crop production.



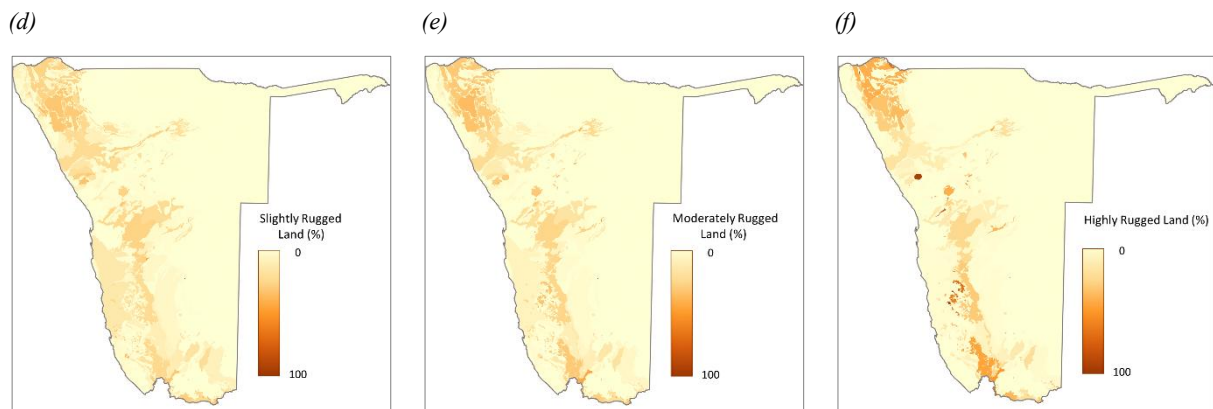


Figure 31. Ruggedness: Percentage of (a) flat; (b) level; (c) nearly level; (d) slightly rugged; (e) moderately rugged; and (f) highly rugged land per soil map unit in Namibia.

The *Topographic Position Index (TPI)* (Figure 32) and the *Topographic Wetness Index (TWI)* are relevant at local scales to identify landscape positions (ridges, upper slopes, footslopes, valley bottoms, and plains) which are correlated with physical, chemical and biological processes acting on the landscape. Erosional and depositional processes influence soil depth, texture, water holding capacity, and accumulation of salts and nutrients.

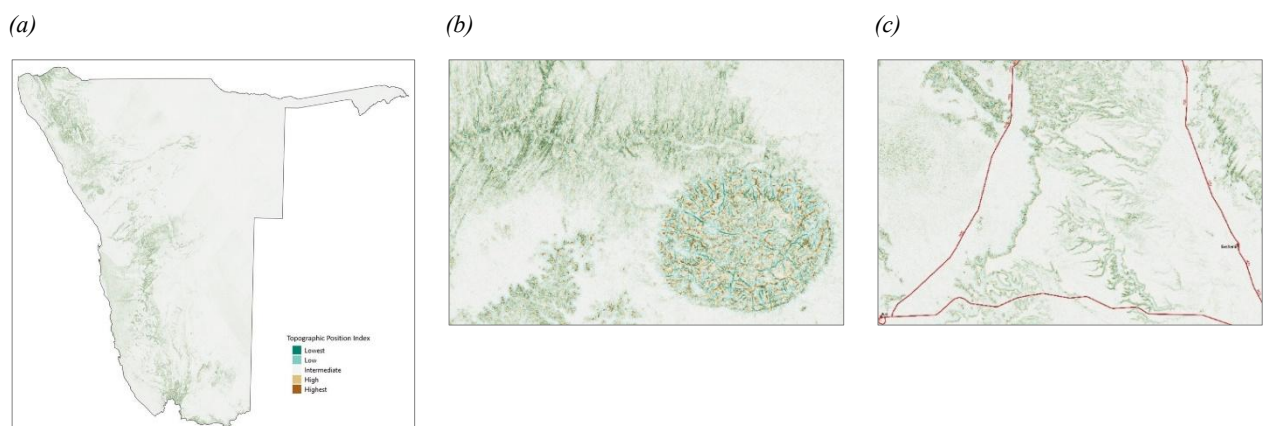


Figure 32. Topographic Position Index (TPI): (a) Namibia; (b) Brandberg–Ugab River area; and (c) Schwarzrand.

5.3.4 Advances and strengths of the national soil map

The NSMN represents a distinct improvement over the Preliminary Soil Map of Namibia produced during the National Soil Survey (NSS) (Coetzee, 2001; ICC & MAWRD, 2000). The most obvious improvement is in spatial detail: 4787 polygons, compared with 321 polygons in the NSS 1M map.

Secondly, the current study has access to better technology and techniques. In 1998–2000, map units were delineated manually on paper prints of Landsat-5 TM imagery and digitised using tablet digitising techniques. This study used a 30 m-resolution DEM, a suite of 30–90 m-resolution terrain derivatives, high-resolution satellite imagery, and the editing capabilities of a modern GIS, with on-screen digitising carried out at working scales between 1:20 000 and 1:50 000.

A third improvement is typological: NSMN distinguishes a substantially larger number and wider diversity of soil topological units (212 versus 36) and soil map units (490 versus 89), capturing more of the compositional complexity evident in Namibian landscapes. By distinguishing

dominant, associated and one or two included soils within each soil map unit, the NSMN represents within-polygon variability more realistically than the NSS 1M, which generally listed a single dominant soil, or complexes or associations of two to three soil types per polygon. The inclusion of multiple principal- and supplementary WRB qualifiers strengthens the diagnostic value of the map, as these qualifiers communicate key functional properties (e.g. shallow depth, carbonates, gypsum, salinity, and texture). The NSS 1M soil typological units typically listed only one or two descriptive adjectives. In the GIS, the map's attribute table lists several practical land-management indicators (e.g. stoniness and depth limitations), increasing the value of the map for agricultural and environmental applications.

A fourth distinction lies in the alignment of the national legend with the 2015 revision of the WRB. This enhances national and international consistency and facilitates integration of Namibian soil information with contemporary regional and global soil datasets.

A fifth difference is technical and practical: the NSMN is a fully digital GIS product integrated with a broad set of ancillary spatial layers, including terrain, landforms, geology, vegetation, satellite imagery, and the NSPD soil profiles and analytical data. This makes the map immediately usable for national planning, environmental assessment and land evaluation, and enables incremental revision as new soil observations become available. In this sense, NSMN is not a one-time, static cartographic product but can evolve over time — an important consideration where data density varies markedly across regions.

Lastly, the NSMN offers spatial coherence and completeness. Previously unmapped areas — the Skeleton Coast and Tsau ǁKhaeb National Parks — were mapped for the first time at a scale larger than 1:4 million. Separate legacy maps of Etosha National Park, the Okavango River terrace, central Owambo, and other localised surveys were spatially and typologically integrated into a single national product. Although mapping confidence remains spatially variable due to uneven legacy map coverage and profile density, the achievement of full national coverage provides a consistent baseline that did not previously exist.

5.3.5 Limitations and sources of uncertainty

Confidence in the NSMN is tempered by the uneven density and spatial distribution of soil observations and the nature of the legacy source maps. Large areas of Namibia have low profile density, and mapping in such areas relied more heavily on interpretation of remote sensing, terrain derivatives, and geological information. On the other hand, regions with more detailed legacy mapping and greater survey effort (e.g. parts of the northeast and north-central regions) allowed greater typological resolution and more defensible boundary placement.

A major methodological challenge in any national synthesis is the integration of products generated at different scales and for different purposes. Although a scale-consistent generalisation strategy was followed, the transition between areas informed by detailed mapping and those derived from reconnaissance-level inputs inevitably introduces variability in polygon size, boundary detail and typological richness. This must be acknowledged to prevent over-interpretation of fine boundary detail in regions where supporting evidence is sparse. The national map can be used with high confidence for applications at 1:250 000 scale or smaller, and moderate confidence at scales between 1:100 000 and 1:250 000. Although the GIS allows more detailed visualisation of areas of interest, it should be noted that zooming beyond the intended usage scale does not imply greater positional or thematic accuracy.

Some soil types and diagnostic conditions fall below the minimum mapping unit at national scale and are therefore only represented as inclusions. This is particularly relevant for spatially restricted soils such as Histosols, Planosols, Luvisols, and Vertisols. In practical terms, this means that NSMN is suitable for national and regional applications but not for local-scale field assessment in high-stakes land-use decisions.

5.4 Digital soil property maps of Namibia

Digital soil mapping forms the predictive component of the NamSIS framework, integrating harmonised soil observations with environmental covariates to generate continuous spatial predictions of soil properties. DSM provides quantitative estimates of soil variability and associated uncertainties across the country. This section presents details of the inputs (soil data and covariates), modelling parameters, performance, outputs (predictions and expressions of uncertainty), pedological evaluation, and external validation.

5.4.1 Soil data

The harmonised national soil dataset brought together measurements collected over several decades by multiple projects, each with its own objectives, sampling depths and methods, and laboratory methods. This produced a dataset typical of legacy soil information systems worldwide: spatially uneven (Appendix I), temporally clustered (Appendix J), and biased towards more accessible and agriculturally important regions. Sampling density is highest across the crop-producing northern regions and lowest in the Namib Desert, national parks, rugged escarpment, and remote communal rangelands such as the northern Kunene Region (former Kaokoveld) and eastern Otjozondjupa Region (former Hereroland East and Bushmanland). The number of observations decline with depth for nearly all properties, reflecting the agricultural bias towards the upper soil layers and practical difficulties of deep soil sampling in the dry, hard, and stony soils of arid environments.

Depth-harmonised descriptive statistics (Table 5) show pronounced skewness for many soil properties, with most values clustered near the lower end of their ranges (Appendices K and L). Sand and BS, on the other hand, frequently approach or reach 100%, consistent with Namibia's widespread sandy deposits and limited leaching.

OC, total N, P, Ca, Mg, K and silt generally decline with depth. This pattern is consistent with minimal leaching under arid conditions and with inputs dominated by surface litter and aeolian dust. Increases in pH and clay content can be ascribed to subsoil carbonate accumulation and eluviation of fine particles — well-documented pedogenic processes in arid and semi-arid regions. With the exception of a few outliers from fertilised crop fields, P and N concentrations are low, which is typical of low-productivity dryland systems (Brady & Weil, 2016). Low CEC and OC reflect the dominance of 1:1 clay minerals and sesquioxides, the sparse vegetation and litter inputs characteristic of drylands, and rapid mineralisation under high soil temperatures.

BS values are generally high or fully saturated. Low rainfall, high evapotranspiration and widespread calcareous and dolomitic lithologies promote base-rich conditions, buffer acidity, and support the formation of shallow calcrete. The Kalahari Basin is dominated by deep, slightly acidic, quartz-rich sands, while the Escarpment and Central Highlands contain finer textures

derived from basaltic, calcareous, dolomitic and schistose parent materials. Topsoil pH (H₂O) span an unusually wide range — from as low as 4 in parts of the Kalahari to more than 11 in arid central and western regions. The high variance in soil properties and weak horizon differentiation across many profiles reflect Namibia's climatic, topographic and lithological heterogeneity.

Table 5. Descriptive statistics of depth-harmonised soil properties across Namibia for three standard depth intervals (0–30, 30–60, and 60–100 cm).

	Bulk Density (kg dm ⁻³)			Base Saturation (%)		
	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm
n	798	616	565	1 786	1 474	1 302
Min	0.75	0.73	1.17	2.12	1.17	1.17
Max	2.01	1.88	2.12	100.00	100.00	100.00
Mean	1.50	1.51	1.52	82.77	81.84	80.74
Median	1.52	1.52	1.51	100.00	100.00	100.00
SD	0.16	0.14	0.15	24.69	28.10	29.77
Skewness	-0.70	-0.60	0.36	-1.31	-1.36	-1.29

	pH (H ₂ O)			Cation Exchange Capacity (cmol(+) kg ⁻¹)		
	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm
n	1 797	1 485	1 313	373	332	266
Min	4.00	4.00	4.00	0.11	0.04	0.06
Max	11.00	11.00	11.00	54.08	54.08	43.10
Mean	7.38	7.49	7.54	6.04	6.15	5.67
Median	7.47	7.73	7.84	3.14	3.82	3.64
SD	1.24	1.46	1.57	7.78	7.08	6.56
Skewness	-0.17	-0.26	-0.26	2.53	2.46	2.59

	Calcium (mg kg ⁻¹)			Magnesium (mg kg ⁻¹)		
	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm
n	721	390	300	720	390	300
Min	1.05	1.55	3.11	0.00	0.00	0.00
Max	9 815.93	6 882.31	7 300.61	761.81	770.87	573.35
Mean	1083.53	744.83	656.77	104.74	105.12	91.82
Median	447.17	347.52	257.49	77.97	66.68	47.23
SD	1 547.25	1 096.64	963.20	99.49	121.29	107.08
Skewness	2.26	2.84	2.89	2.20	2.13	1.71

	Potassium (mg kg ⁻¹)			Sodium (mg kg ⁻¹)		
	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm
n	722	390	300	721	390	300
Min	0.01	0.01	0.38	0.00	0.00	0.00
Max	942.12	908.24	894.59	880.23	1684.00	833.44
Mean	120.29	120.05	97.22	43.60	52.39	44.29
Median	94.00	90.27	60.24	17.10	21.75	18.17
SD	109.64	121.04	120.48	110.08	146.29	93.48
Skewness	2.22	2.06	3.04	5.29	7.64	5.07

	Electrical Conductivity (1:2.5) (μS cm ⁻¹)			Electrical Conductivity (sat) (μS cm ⁻¹)		
	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm
n	1 553	1 216	863	580	547	507
Min	4.12	2.00	2.00	2.42	2.08	2.22
Max	11 775.24	12 747.96	8 039.04	10 121.06	13 214.33	10 359.98
Mean	211.61	257.43	260.20	604.33	1 177.64	1 335.95
Median	50.67	44.87	46.46	104.50	147.37	185.58
SD	798.90	910.89	835.14	1 337.86	2 035.19	1 995.57
Skewness	8.77	7.14	5.60	3.97	2.48	1.95

	Organic Carbon (%)			Phosphorus (mg kg ⁻¹)		
	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm
n	2 091	1 573	1 252	726	623	507
Min	0.00	0.00	0.00	0.01	0.01	0.01
Max	3.11	2.76	3.26	31.19	44.97	49.00
Mean	0.35	0.27	0.22	3.55	2.17	1.73
Median	0.28	0.20	0.15	1.99	0.65	0.47
SD	0.29	0.24	0.26	4.23	3.96	4.13
Skewness	3.35	3.82	5.10	2.27	4.27	6.69

	Nitrogen (mg kg ⁻¹)			Sand (%)		
	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm
n	572	330	157	1 937	1 411	1 030
Min	30.00	10.00	10.00	11.00	2.10	2.10
Max	2 500.00	1 284.82	716.74	100.00	99.95	100.00
Mean	413.06	322.24	188.99	79.56	76.85	77.61
Median	308.21	243.61	160.93	83.80	80.91	81.92
SD	324.41	263.04	118.68	15.45	16.84	17.20
Skewness	2.34	1.93	1.65	-1.26	-1.11	-1.19

	Silt (%)			Clay (%)		
	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm
n	1 681	1 209	1 030	1 681	1 209	1 030
Min	0.00	0.00	0.00	0.00	0.02	0.00
Max	65.71	73.96	74.30	56.55	67.59	70.51
Mean	9.55	9.51	9.02	9.61	12.00	13.11
Median	6.50	6.08	5.12	7.00	9.00	10.25
SD	9.14	9.04	9.33	8.36	10.28	11.08
Skewness	1.72	1.77	1.77	2.04	1.92	1.71

5.4.2 Feature selection

The *Boruta* feature selection algorithm, supplemented by manual adjustments where necessary, reduced the covariate set per model to between 4 and 60 predictors (Appendix M), retaining only those variables that contributed meaningfully to model performance. Covariate selection frequencies are summarised in Table 6.

The DEM, rainfall, temperature, potential evapotranspiration, downstream flow length, presence of rock outcrops and vegetation indices were frequently selected. In contrast, convexity and the aridity index were never selected, and variables such as aspect, eastness, northness, curvature metrics, flow direction, flow accumulation and distance-to-drainage were selected in fewer than 10% of models. These low-frequency predictors contribute only weakly at national scale, but they may still be locally relevant.

ASTER-derived mineralogical indices (kaolinite, quartz, calcite, carbonates and mafic minerals) were initially tested and often contributing meaningfully but were largely excluded from final models due to visible mosaicking artefacts.

A notable finding is the inclusion of locally produced conventional mapping products — vegetation types, soil classes, geology, agro-ecological zones, and carrying capacity — in more than half of the models.

Table 6. Frequency (%) with which environmental covariates were included in all Random Forests models.

Covariate	Freq	Covariate	Freq	Covariate	Freq	Covariate	Freq
DEM	94	SAVI _w	65	NIR _w	52	ALOS landform	15
Precipitation	90	KNDVI _w	63	SWIR2 _s	52	Eastness	8
Temperature	88	NIR _s	63	SWIR1 _s	50	Flow accumulation	8
ASPP	79	AEZ (2021)	60	TPI	50	Distance to drainage	8
Flow length down	79	Carbonate norm	60	SWIR1 _w	48	Maximum curvature	6
EVI _w	73	Green _s	60	Iron norm	46	Flow direction	4
MSAVI _s	73	NDVI _w	60	Clay norm	44	Horizontal curvature	4
Vegetation types	73	ALOS topo divers	60	Land cover	42	Northness	4
MSAVI _s	71	Blue _s	58	Iwahasi landform	42	Aspect	2
SAVI _s	71	KNDVI _s	58	Red _w	40	Carbonate index	2
AEZ (1996)	69	NDVI _s	58	Ferrous min norm	38	Kaolinite index	2
Blue _w	69	Namibia soil map	56	Slope	35	GAI	0
EVI _s	69	Geology (2022)	54	Height above drainage	25	Calcite index	0
PET	69	Green _w	54	Vertical curvature	23	Convexity	0
Rock outcrops norm	69	Red _s	54	Flow length upstream	15	Mafic index	0
Carrying capacity	67	SWIR2 _w	54	Lithology (Geology 1977)	15	Quartz index	0
CHILI	65						

Note: The _w and _s suffixes associated with spectral bands and vegetation indices refer to two seasonal windows: winter and summer. More covariate information and clarification on the abbreviations/acronyms are provided in Appendix B.

5.4.3 Model optimisation

Model-based optimisation of RF hyperparameters, carried out with the *tuneRanger* R package, increased predictive accuracy and generalisability. Guided by DSM literature (Hengl et al., 2017b; Lopes, 2019; Zhang et al., 2017) and model-stability tests, the number of trees (*ntree*) was fixed at 150, beyond which model performance plateaued. *Max nodes* was left unconstrained (*NULL*), allowing trees to grow to their natural maximum size and minimise bias, at the potential cost of increased variance. Other hyperparameters — *nodesize*, *mtry*, and *samsize* — were tuned independently for each soil property–depth combination (Appendix N). This property-specific optimisation was essential given the broad spectrum of soil processes represented and the differing degrees of spatial autocorrelation, lithological control, and environmental predictability across variables.

5.4.4 Model performance (quantitative results)

Model performance varies among soil properties and depth intervals. Table 7 summarises the mean model performance metrics across 20 bootstrap iterations, while descriptive statistics of all bootstrap realisations are provided in Appendix O.

Properties that are strongly influenced by climate, parent material, and vegetation productivity tend to be predicted more reliably. *pH*, *BS*, and *extractable Ca* and *K* are closely association with lithological controls, carbonate presence, and the marked gradients in rainfall and evapotranspiration across Namibia. Particle-size fractions also performed relatively well, particularly *sand* and *silt* content, due to the strong regional influence of aeolian deposits, Kalahari sediments, alluvial systems, and weathering environments on soil texture distribution across Namibia.

Table 7. Mean model performance metrics (RMSE and R^2) across 20 bootstrap iterations for each soil property and depth interval.

	R^2			RMSE		
	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm
Bulk density (kg dm^{-3})	0.37	0.21	0.29	0.13	0.12	0.13
Base saturation (%)	0.45	0.44	0.46	18.3	21.3	21.7
pH (H_2O)	0.52	0.46	0.49	0.87	1.07	1.13
Cation exchange capacity ($\text{cmol}(+) \text{kg}^{-1}$)	0.46	0.49	0.19	5.5	5.1	6.0
Calcium (mg kg^{-1})	0.46	0.45	0.32	1081	830	836
Magnesium (mg kg^{-1})	0.31	0.27	0.30	82.2	101.8	90.3
Potassium (mg kg^{-1})	0.40	0.42	0.32	87.0	91.9	96.0
Sodium (mg kg^{-1})	0.11	0.15	0.22	95.7	114.2	76.9
Electrical conductivity (1:2.5) ($\mu\text{S cm}^{-1}$)	0.35	0.19	0.10	614	792	754
Electrical conductivity (sat) ($\mu\text{S cm}^{-1}$)	0.13	0.37	0.39	1266	1602	1512
Organic carbon (%)	0.31	0.18	0.15	0.24	0.21	0.22
Phosphorus (mg kg^{-1})	0.31	0.27	0.19	3.59	3.50	3.74
Nitrogen (mg kg^{-1})	0.57	0.75	0.25	211.3	131.7	101.5
Sand (%)	0.45	0.36	0.33	11.6	13.3	13.9
Silt (%)	0.44	0.39	0.34	7.0	7.2	7.5
Clay (%)	0.32	0.22	0.28	7.0	9.2	9.1

Properties that are strongly influenced by climate, parent material, and vegetation productivity tend to be predicted more reliably. *pH*, *BS*, and *extractable Ca* and *K* are closely association with lithological controls, carbonate presence, and the marked gradients in rainfall and evapotranspiration across Namibia. Particle-size fractions also performed relatively well, particularly *sand* and *silt* content, due to the strong regional influence of aeolian deposits, Kalahari sediments, alluvial systems, and weathering environments on soil texture distribution across Namibia.

The relatively good performances of *CEC* and *N* models are also consistent with their dependence on texture, mineralogy, and organic matter inputs, the latter indirectly captured through climatic and spectral covariates, especially vegetation indices.

OC and *P* are influenced by local biological cycling and land management history, which are difficult to capture using broad-scale environmental covariates. Analytical inconsistencies among legacy laboratory datasets likely contributed additional uncertainty for these properties. The relatively good RMSE values indicate that the models reproduce realistic spatial trends despite modest coefficients of determination.

The weak performance of *Na* and *EC* models can be attributed to both pedological and dataset limitations. Salt-affected soils are spatially discontinuous and often associated with highly localised hydrological conditions such as pans, ephemeral drainage lines, and shallow groundwater systems. Agricultural surveys avoided known saline soils, while extreme values from a small number of salt pan samples introduced strong skewness into the training data, reducing model stability.

Across all soil properties, predictive accuracy is higher for topsoil than subsoil. Topsoil predictions benefited from denser observations, stronger coupling with remotely sensed environmental variables, and more direct relationships with vegetation and surface hydrological processes. Subsoil properties are less directly linked to surface covariates and are additionally affected by

incomplete profile descriptions and variable sampling depths. The depth-related declines in model performance represent a widely reported, persistent challenge in national-scale modelling (e.g. Hengl et al., 2017a; Liu et al., 2022b).

Overall, the performances of the models are within the ranges reported in other national-scale DSM studies (Appendix T) employing single-model approaches, particularly in regions characterised by environmental heterogeneity and limited data availability, although somewhat lower than results obtained using more advanced ensemble or hybrid modelling techniques. However, comparison with other studies should be made with caution, given differences in environmental settings, data variance, depth intervals, laboratory methods, and modelling algorithms.

5.4.5 Variable importance and environmental controls on pedogenesis

An analysis of the variable importance patterns (Figure 33) provides insight into the environmental controls governing soil variability across Namibia.

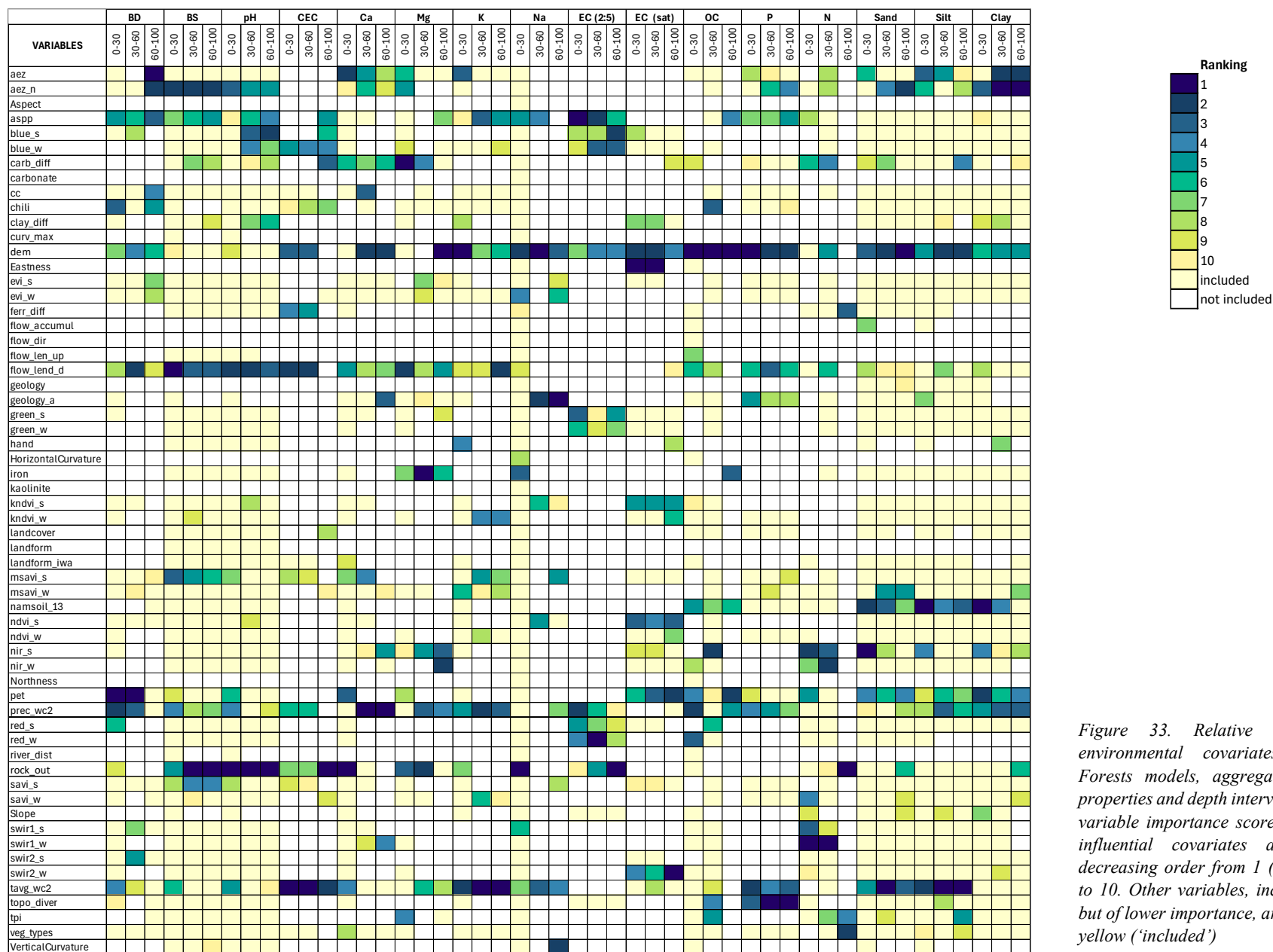
Climatic variables were among the most influential predictors across many models, underscoring the fundamental role of water availability in Namibian soil formation. Gradients in rainfall, aridity, temperature, and evapotranspiration affect weathering intensity, leaching, organic matter input and turnover, carbonate accumulation, salinity development, and vegetation productivity. The broad transition from hyper-arid western Namibia to semi-arid northeastern savannas is mirrored by major shifts in soil properties and Reference Soil Groups. These patterns suggest that climatic controls provide much of the large-scale structure visible in both the conventional and digital mapping outputs.

Terrain variables contributed substantially to several models, as elevation, slope, terrain ruggedness, and hydrological indices influence redistribution of sediments, runoff concentration, erosion, deposition, and drainage conditions. The influence of terrain metrics was generally stronger at broader physiographic scales, while micro-topographic metrics that represent fine-scale hillslope processes were of less importance. This likely reflects both the predominance of extensive low-relief landscapes across much of Namibia and the 90 m modelling resolution.

Vegetation indices and *land-cover* variables contributed meaningfully to several soil-property models, particularly OC and N. Vegetation acts both as a response to environmental gradients and as an active soil-forming factor through organic matter inputs, root activity, nutrient cycling, and effects on soil moisture. In arid environments, vegetation patterns often closely reflect water availability and substrate conditions, explaining the predictive value of remotely sensed vegetation indicators in DSM.

The importance of *parent material* proxies, including geology and conventional soil mapping units, highlights lithological control on weathering pathways, particle-size distribution, salinity, carbonate accumulation, and mineral composition in Namibia.

The contributions of the locally produced categorical layers (soil, vegetation, agro-ecological zones, and carrying-capacity maps) is particularly significant because these datasets contain contextual landscape information and embedded expert insight into environmental processes and relationships that are difficult to capture through purely statistical relationships using global raster covariates alone.



With the notable exception of the Na models, the spatial patterns confirm that the DSM models are capturing major pedogenic processes. The mechanisms underlying predictor importance vary among individual soil properties:

BD depends strongly on vegetation-related covariates and climate, showing the influence of soil organic matter (SOM) on *BD* and compressibility (Rühlmann & Körschens, 2009). Soil desiccation can cause shrinkage of the soil matrix, cracking and structural rearrangement, especially of smectite-rich soils (Tang et al., 2021), which could potentially account for the influence of the ‘distance to drainage lines’ covariate.

BS, *Ca*, *Mg*, and *K* are significantly influenced by proxies for vegetation cover (primary production, vegetation indices), rainfall, site drainage, and the type and depth of rocks (particularly shallow and exposed carbonate-rich substrates) through the supply of bases, buffering of soil acidity, and limited deep percolation and leaching.

The importance of near-infrared (NIR) and shortwave-infrared (SWIR) covariates for *N* probably reflects indirect relationships with vegetation productivity, canopy water content, surface moisture and soil organic matter (Mulder et al., 2011).

OC is linked to the DEM, vegetation indices and spectral bands (especially red and NIR), which is consistent with moisture and temperature regulation of SOM accumulation and mineralisation (Angelopoulou et al., 2019; Mulder et al. 2011).

Texture also relates to the DEM, reflecting the typical separation of coarse and fine materials along slopes, where erosion removes finer particles from upper positions and deposits them downslope. The strong contribution of the local soil and agro-ecological zone maps highlights the major geomorphic and lithological contrasts across Namibia — particularly between the coarse quartz-rich Kalahari sands and the finer-textured soils of the Central Highlands, formed from basaltic, calcareous, dolomitic, and schistose parent materials. Climatic predictors such as precipitation, temperature, and evapotranspiration are also influential in the sand, silt, and clay models. These factors control the soil moisture regime, which in turn affects weathering intensity, the formation of clay minerals, and the degree to which wind and water can selectively remove fine particles from the topsoil. Together, these processes explain why climate variables are key determinants of the modelled particle size fractions.

5.4.6 Spatial predictions and pedological plausibility assessment

Spatially continuous 90 m resolution maps were generated for all properties and three depth classes by averaging the 20 bootstrap outputs (Figure 34–Figure 41).

A prominent diagonal transition is evident in many soil property maps: the sharp contrast between the diverse geology of the Escarpment and Central Highlands and the deep, quartz-rich, relatively homogeneous sands of the Kalahari Basin. This is reinforced by the southwest–northeast rainfall gradient, which strongly influences leaching intensity, vegetation productivity, and base mobility.

In the Kalahari, somewhat higher rainfall and low clay content promote the leaching of bases, producing low topsoil pH, *BS*, CEC, *Ca*, *Mg*, *K*, *Na* and *EC*. Conversely, the central, western and southern regions show higher concentrations of bases due to calcareous lithologies, limited leaching and high evapotranspiration.

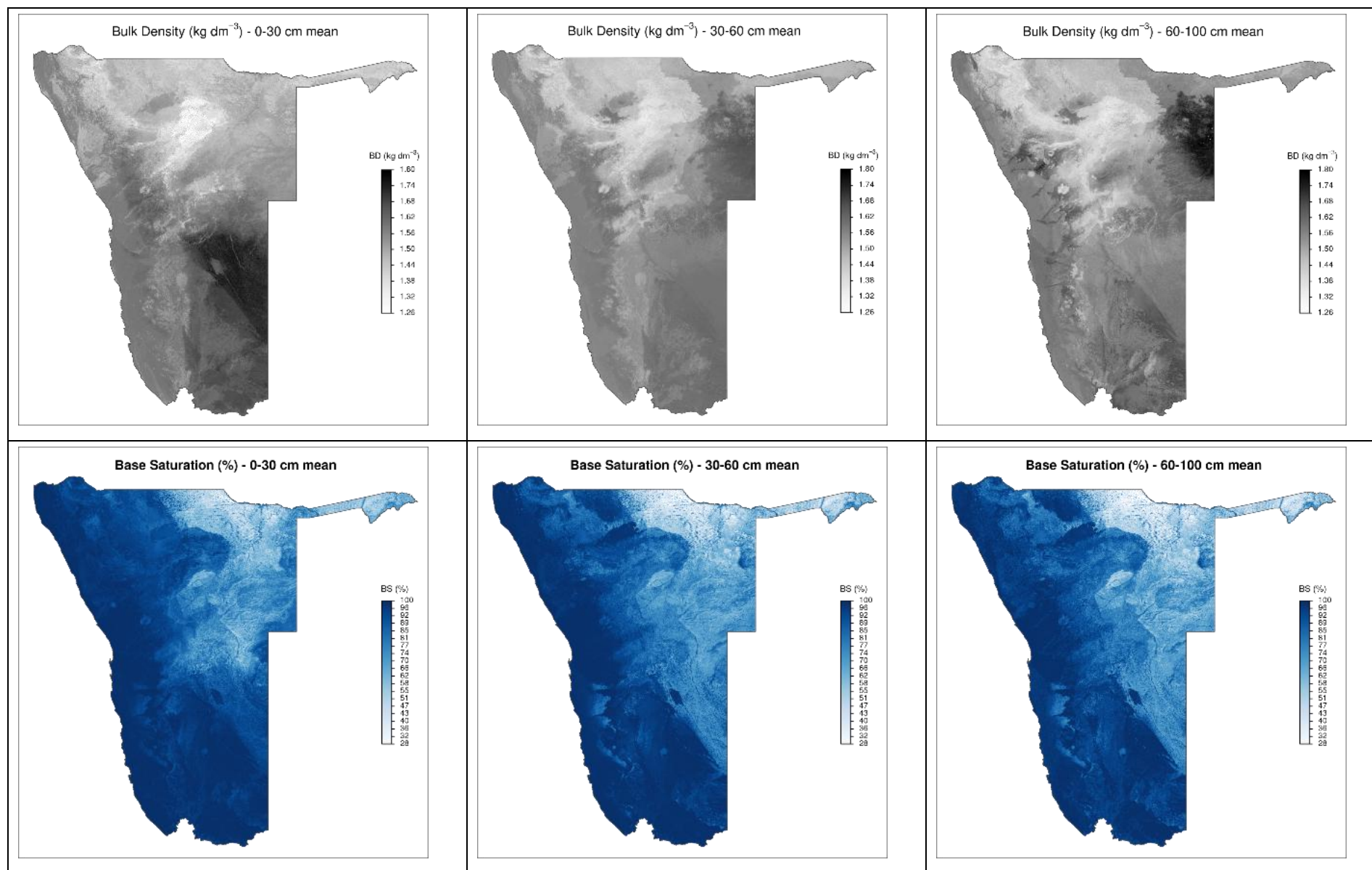


Figure 34. National-scale spatial predictions of bulk density (BD) (top) and base saturation (BS) (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at <https://soilsofnamibia.com/properties.html>).

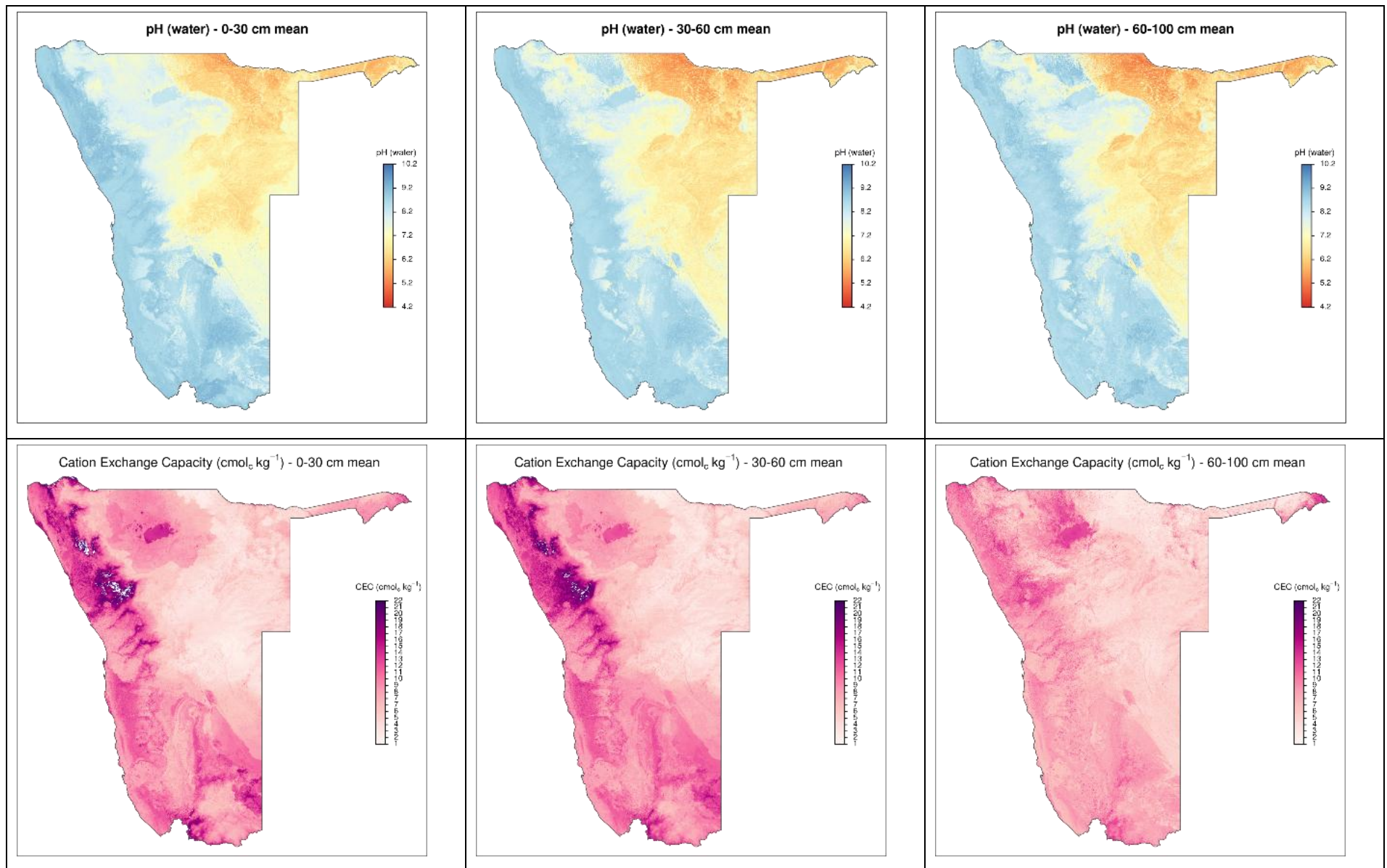


Figure 35. National-scale spatial predictions of pH (H₂O) (top) and cation exchange capacity (CEC) (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at <https://soilsofnamibia.com/properties.html>).

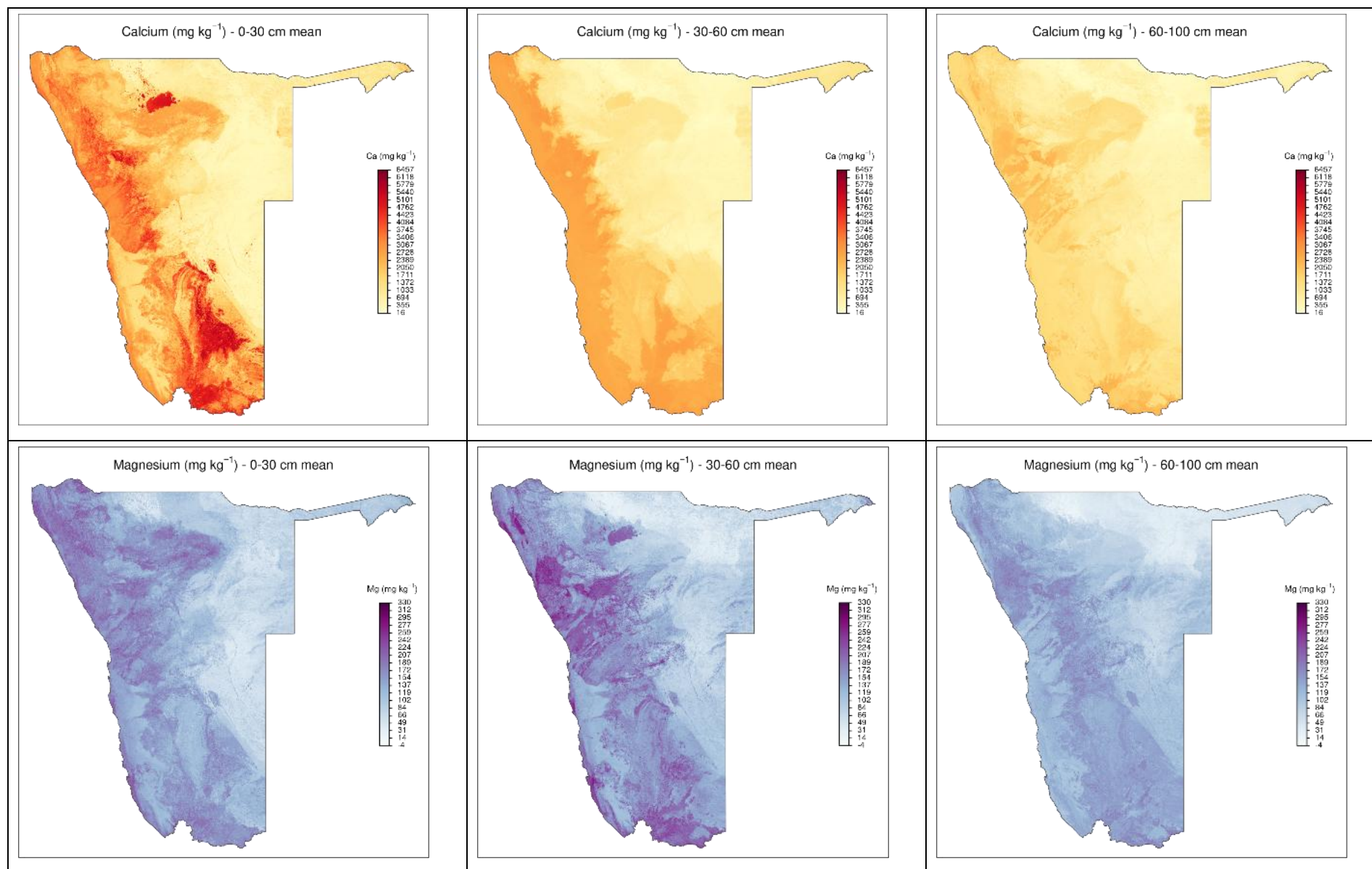


Figure 36. National-scale spatial predictions of extractable calcium (Ca) (top) and magnesium (Mg) (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at <https://soilsofnamibia.com/properties.html>).

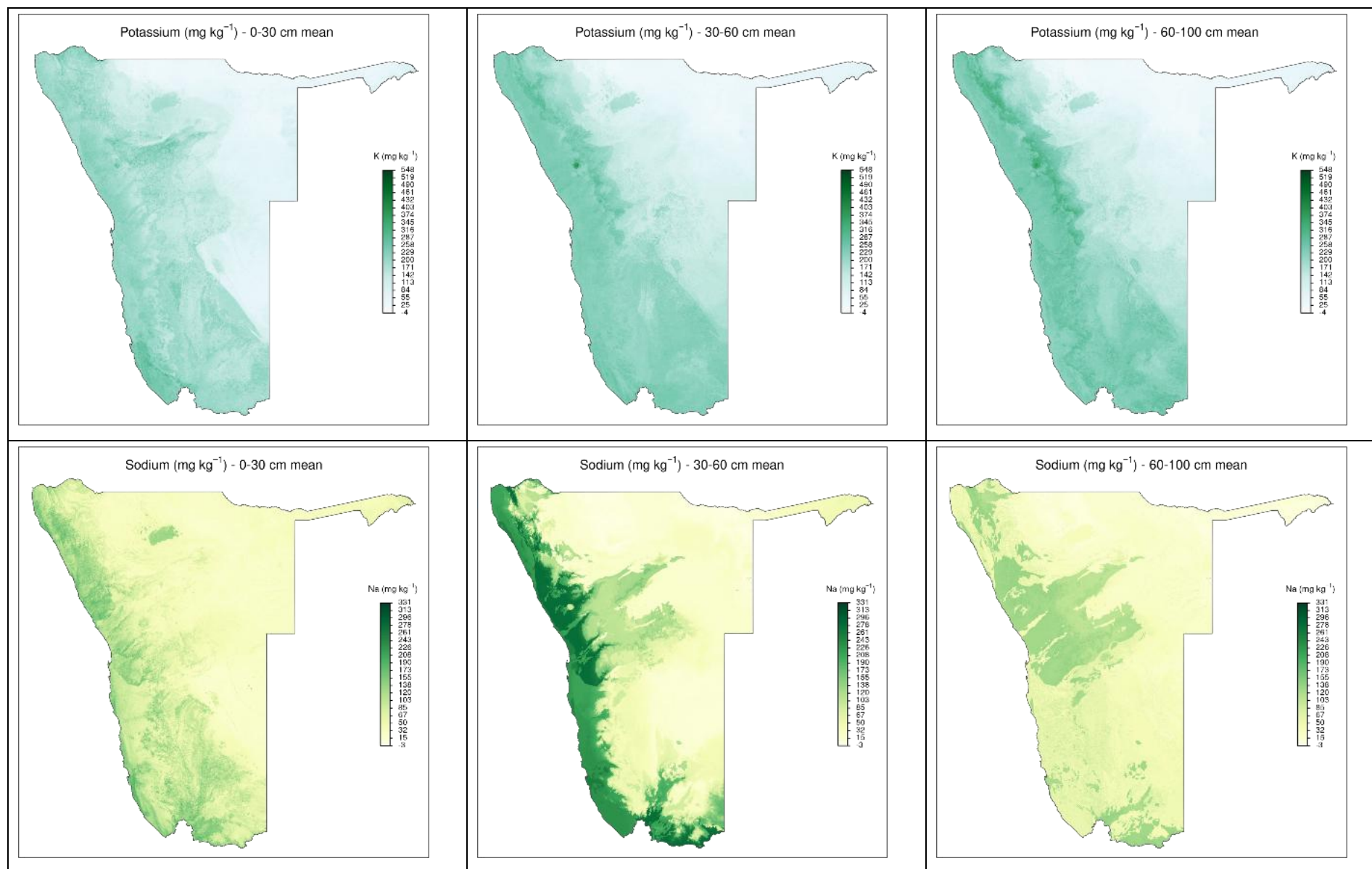


Figure 37. National-scale spatial predictions of extractable potassium (K) (top) and sodium (Na) (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at <https://soilsofnamibia.com/properties.html>).

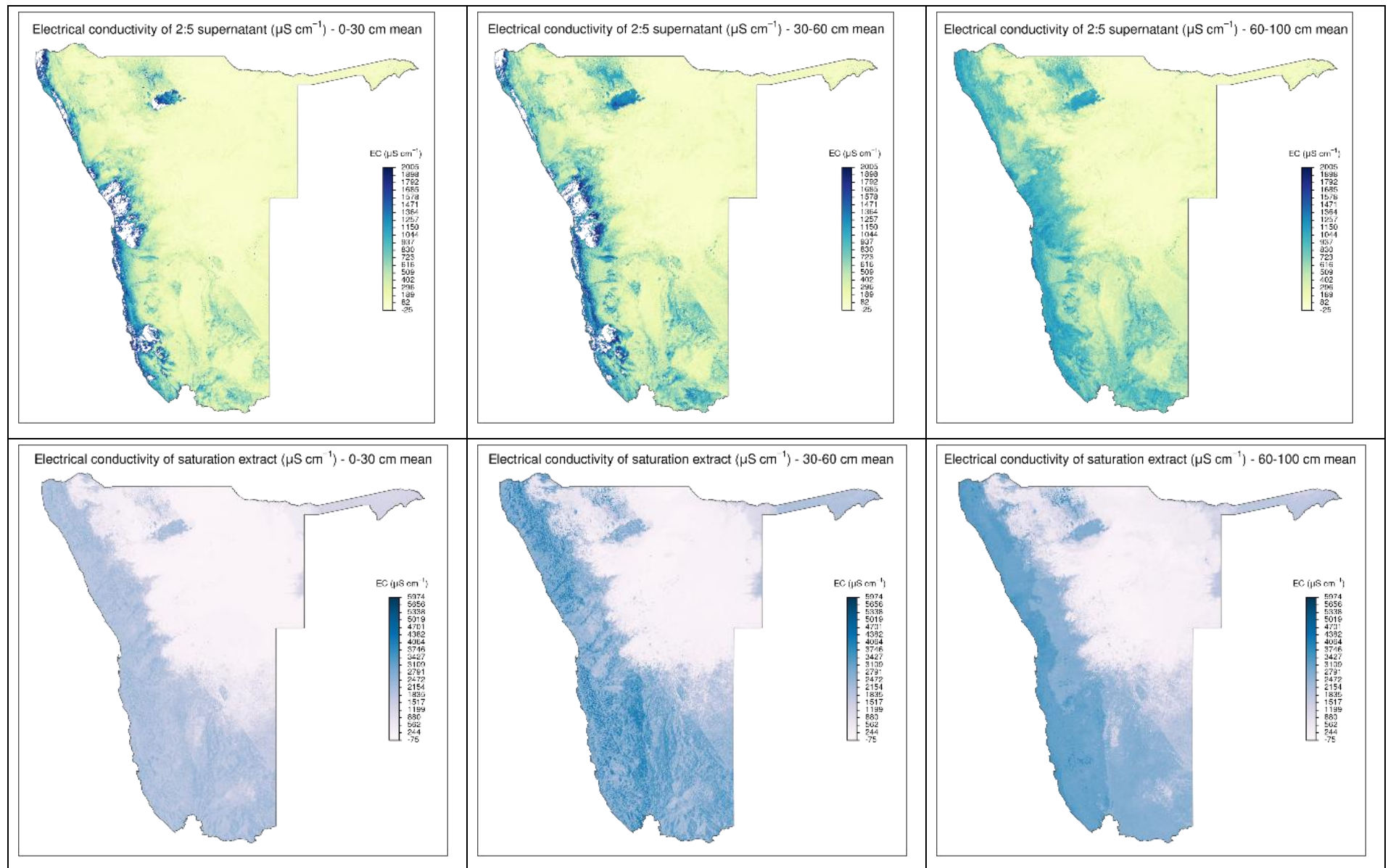


Figure 38. National-scale spatial predictions of electrical conductivity (EC) measured in a 1:2.5 soil:water extract (top) and in a saturated paste extract (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at <https://soilsofnamibia.com/properties.html>).

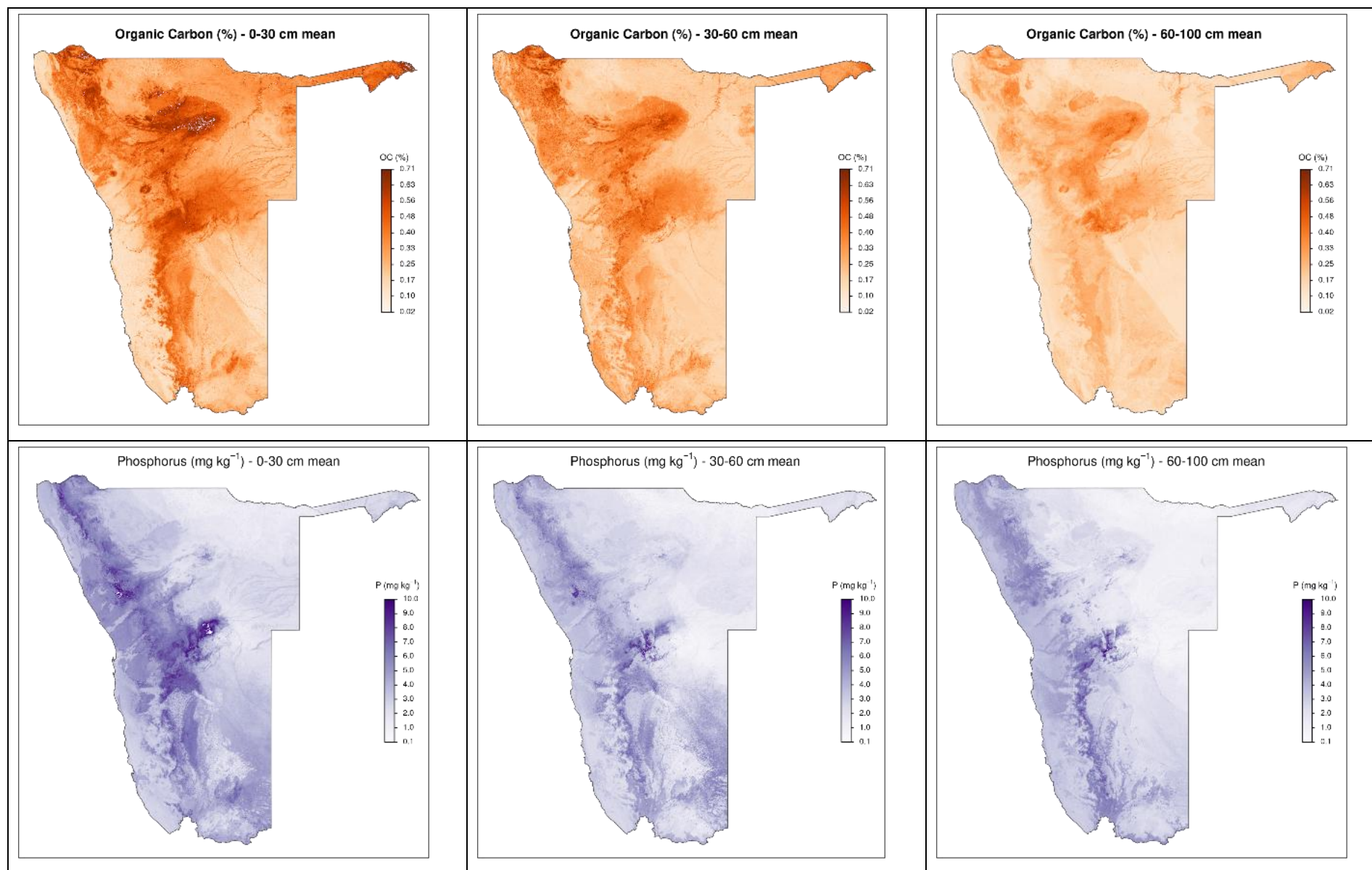


Figure 39. National-scale spatial predictions organic carbon (OC) (top) and plant-available phosphorus (P) (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at <https://soilsofnamibia.com/properties.html>).

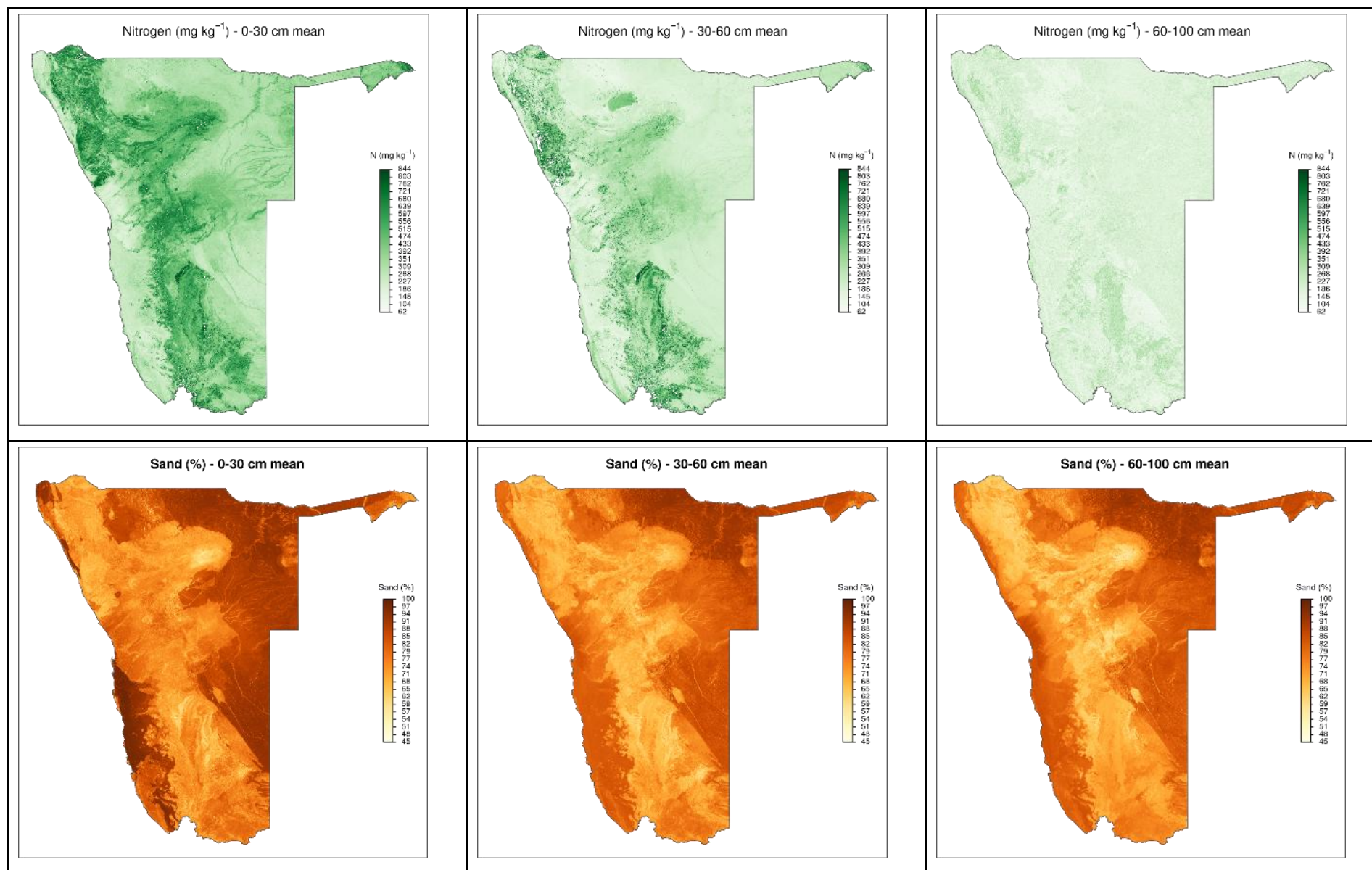


Figure 40. National-scale spatial predictions of total nitrogen (N) (top) and sand content (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at <https://soilsofnamibia.com/properties.html>).

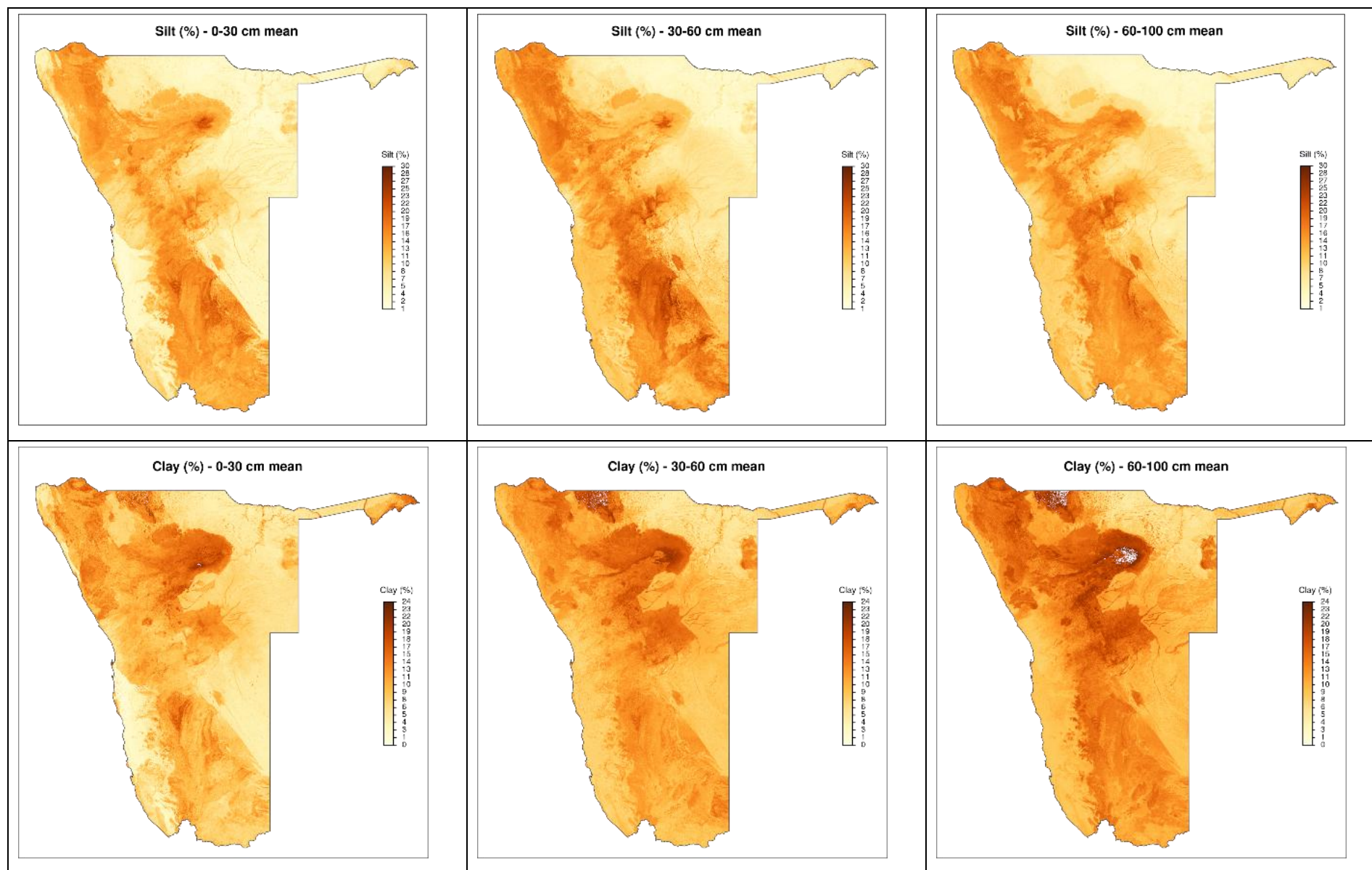


Figure 41. National-scale spatial predictions of silt (top) and clay content (bottom) at 90 m resolution, shown as the mean of 20 bootstrap realisations for three depth intervals (0–30, 30–60, and 60–100 cm) (left to right) in Namibia. (Full resolution maps can be viewed and downloaded at <https://soilsofnamibia.com/properties.html>).

OC and N are highest in the more humid northeast and along ephemeral rivers, where deep-rooted riparian vegetation accesses alluvial aquifers and generates SOM inputs that contrast sharply with adjacent dryland surfaces. P enrichments correspond to limestone–dolomite (e.g. Karstveld), schistose (e.g. Khomas Hochland), and basaltic (e.g. Etendeka Plateau) landscapes.

Texture maps reflect the dominance of sand in the Kalahari and the presence of finer textures in the alluvial floodplains of the Cuvelai, Okavango, Kwando–Linyanti–Chobe and Zambezi rivers, and the Central Highlands and areas underlain by basalt, limestone, dolomite, phyllite and schist. A slight increase in clay and silt at depth can be ascribed to typical eluviation processes in semi-arid landscapes.

The extractable Na models show both low statistical accuracy and unrealistic spatial patterns. This outcome is likely due to sampling bias in the legacy data. Although the results are presented here for completeness, the Na maps should not be disseminated to the public.

The qualitative visual assessment confirmed that predicted spatial distributions are generally consistent with known soil–terrain relationships, climate and vegetation gradients, lithological controls, and dominant pedogenic processes across Namibia, with the notable exception of Na.

5.4.7 Prediction uncertainty

Uncertainty was quantified for each property and depth by mapping the per-pixel 5th and 95th percentiles and the 90% prediction interval (PI90) from the suite of 20 bootstraps, as shown for pH in Figure 42 and Figure 43. The corresponding maps of the other soil properties can be found in Appendix P. As expected, uncertainty increases with depth for most properties and is highest in under-sampled regions (Appendix Q), a pattern also noted by [Hengl et al. \(2014\)](#) and [Poggio et al. \(2021\)](#) in continental DSM products.

Table 8. Prediction Interval Coverage Probability (PICP) for each soil property and depth interval, to assess the robustness of the 90% prediction intervals.

	PICP 90		
	0–30 cm	30–60 cm	60–100 cm
Bulk density (kg dm ⁻³)	92.1	89.4	90.1
Base saturation (%)	90.6	90.6	90.7
pH (H ₂ O)	90.5	91.1	91.0
Cation exchange capacity (cmol(+) kg ⁻¹)	90.3	91.9	94.4
Calcium (mg kg ⁻¹)	91.3	91.0	92.3
Magnesium (mg kg ⁻¹)	94.4	93.8	92.0
Potassium (mg kg ⁻¹)	94.5	92.1	96.0
Sodium (mg kg ⁻¹)	96.7	96.9	96.3
Electrical conductivity (1:2.5) (μS cm ⁻¹)	96.7	95.3	95.6
Electrical conductivity (sat) (μS cm ⁻¹)	94.5	92.3	91.1
Organic carbon (%)	94.7	95.2	96.0
Phosphorus (mg kg ⁻¹)	93.3	95.2	96.6
Nitrogen (mg kg ⁻¹)	94.9	90.0	89.8
Sand (%)	91.3	91.7	91.2
Silt (%)	92.8	92.7	93.0
Clay (%)	93.8	93.3	93.6

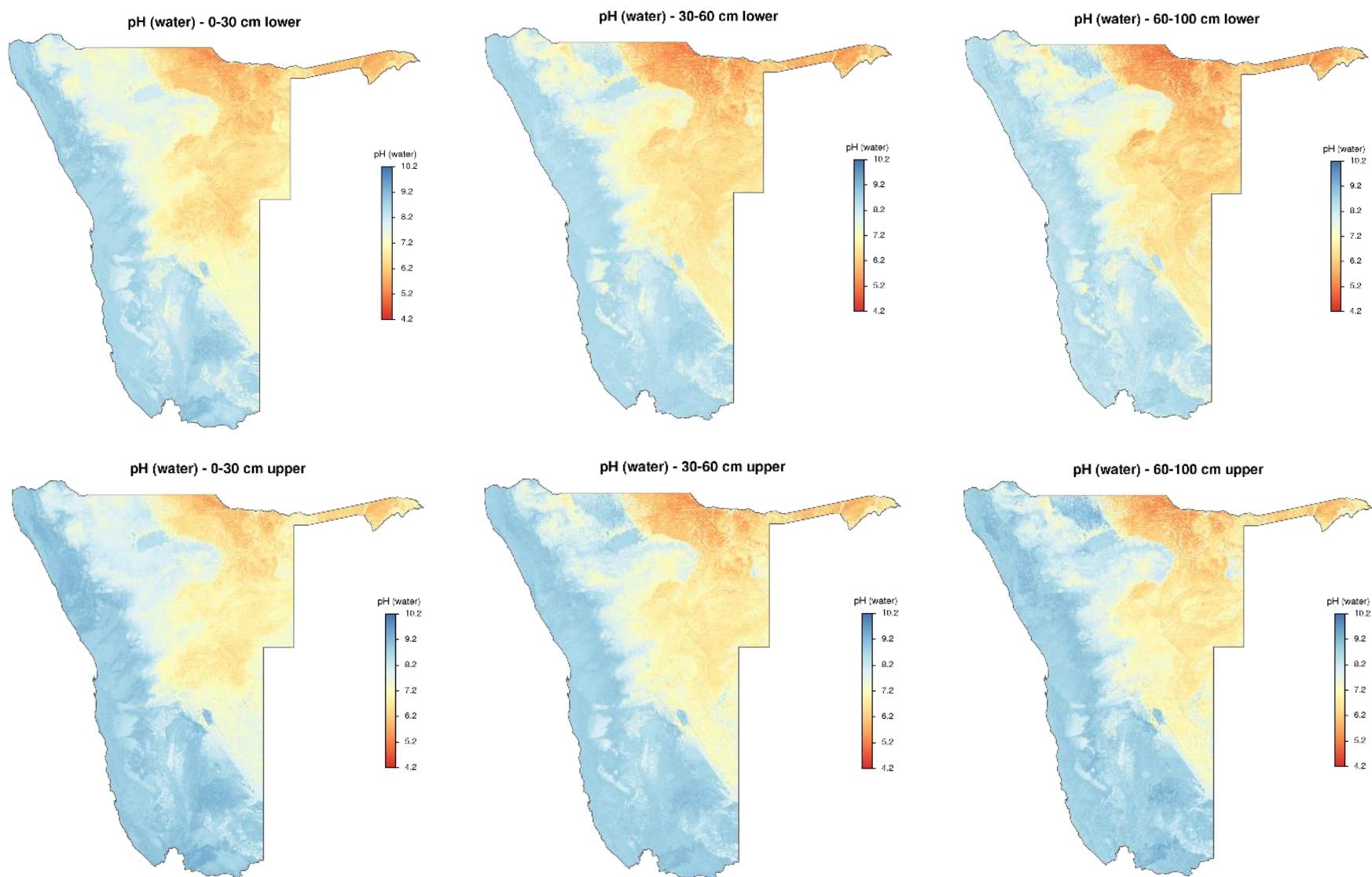


Figure 42. Spatial distribution of prediction uncertainty for pH (H_2O) in Namibia. For each depth interval (0–30, 30–60, and 60–100 cm) (left to right), maps show the 5th (top row) and 95th percentile (bottom row) predictions. (Full resolution maps can be viewed and downloaded at <https://soilsofnamibia.com/properties.html>).

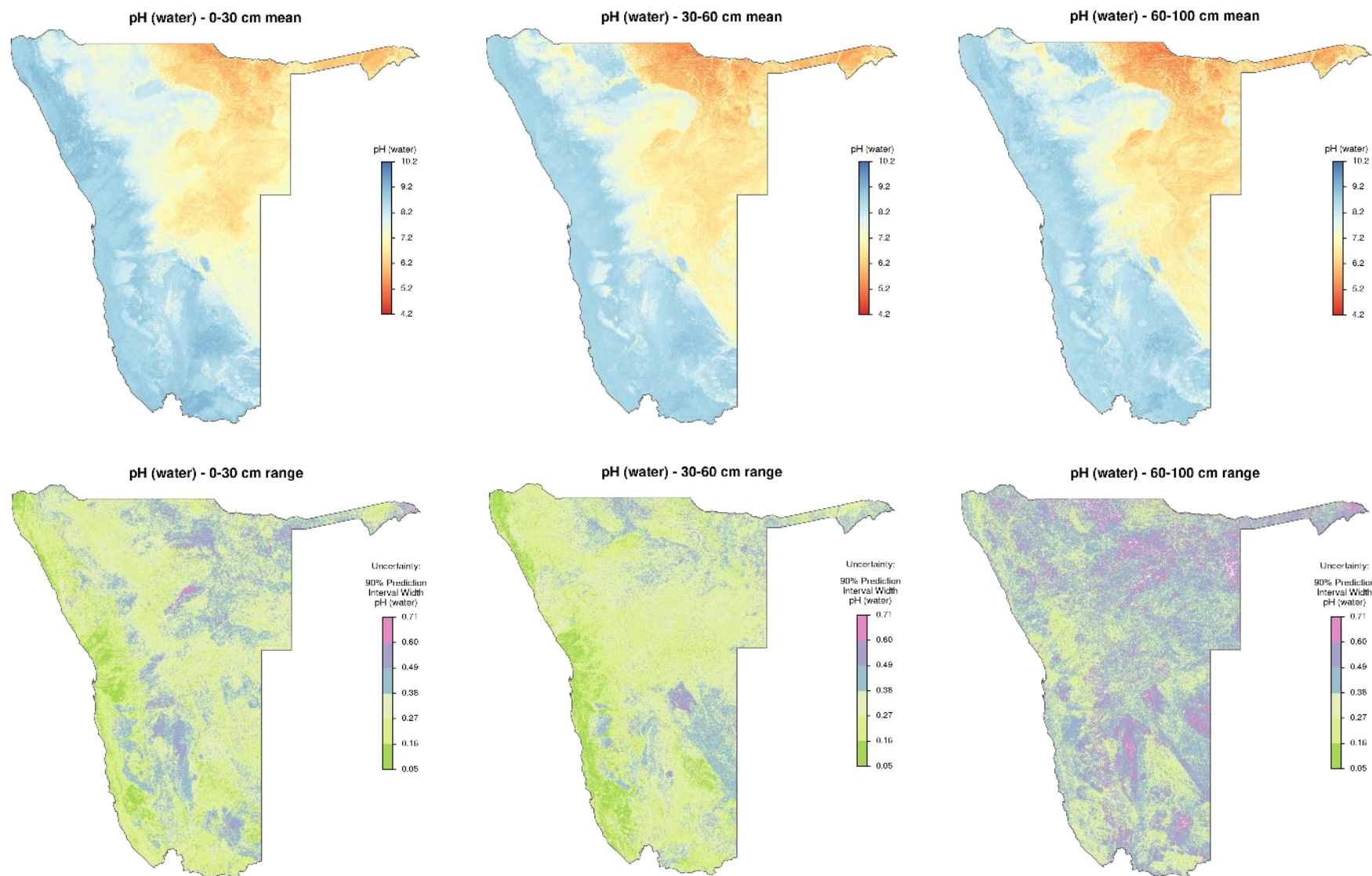


Figure 43. Spatial distribution of prediction uncertainty for pH (H₂O) in Namibia. For each depth interval (0–30, 30–60, and 60–100 cm) (left to right), maps show the mean prediction (top row) and 90% prediction interval (PI90) (bottom row). (Full resolution maps can be viewed and downloaded at <https://soilsofnamibia.com/properties.html>).

The Kalahari exhibits relatively low uncertainty for all properties except BS and pH, reflecting its relatively uniform parent materials and pedogenic processes. More heterogeneous terrains exhibit wider PI90 values.

Prediction Interval Coverage Probability (PICP) results (Table 8; Appendix R) indicate that the pH, BS, Ca, CEC, and sand models achieved values close to the nominal 90% target. Most models overestimate uncertainty to some extent, especially for OC, P, EC and Na. The tendency toward conservative uncertainty estimates likely reflects sparse and uneven sampling, high intrinsic variability in some properties, heterogeneous landscapes that amplify local deviations, the variance-inflating nature of bootstrapping, and model instability in poorly sampled areas. In practical terms, most actual values can be expected to fall within the predicted uncertainty range at any given location.

5.4.8 External validation

Predicted soil property values from this study, SoilGrids250m V2, and iSDA were validated using independent measurements from the Soils4Africa dataset, as summarised in Table 9 and Appendix S.

Table 9. External validation metrics (RMSE and R^2) comparing predictions from this study with iSDA and SoilGrids250m V2 products, against independent Soils4Africa observations.

Variable	Depth (cm)	Namibia	iSDA	SoilGrids	Namibia	iSDA	SoilGrids
		R^2			RMSE		
pH (H ₂ O)	0–20	0.44	0.15	0.37	1.07	1.02	1.22
	20–50	0.35	0.15	0.32	1.03	1.13	1.37
CEC (cmol(+) kg ⁻¹)	0–20	0.16	-	0.07	5.9	-	10.1
	20–50	0.13	-	0.16	4.5	-	8.7
Organic carbon (%)	0–20	0.20	0.17	0.01	0.21	0.20	0.34
	20–50	0.21	0.19	0.08	0.17	0.16	0.18
Nitrogen (mg kg ⁻¹)	0–20	0.61	0.43	0.39	192.4	213.3	333.1
	20–50	0.48	0.37	0.29	210.0	188.9	227.1
Sand (%)	0–20	0.33	0.35	0.14	8.1	16.2	13.3
	20–50	0.34	0.33	0.21	8.5	15.3	12.5
Silt (%)	0–20	0.51	0.39	0.35	3.2	6.4	4.8
	20–50	0.12	0.08	0.07	6.9	7.9	7.6
Clay (%)	0–20	0.06	0.21	0.02	6.6	9.5	10.2
	20–50	0.25	0.31	0.08	6.1	9.2	9.9

Although the Soils4Africa dataset only includes a small number of unevenly distributed Namibian sampling sites, the limited assessment still offers an indication of model performance compared to the other studies. Across all evaluated soil properties, the models produced in this study yielded RMSE values that were consistently comparable to or lower (better) than those obtained from SoilGrids and iSDA. The largest reductions in RMSE were observed for the particle size fractions. Likewise, the coefficients of determination (R^2) of this study were generally higher (better) than those of the two reference products, with the exception of subsoil CEC and topsoil clay. The R^2 of CEC, clay, and OC were poor in all three datasets, probably caused by limitations of the validation dataset rather than model inadequacy. The Soils4Africa samples from Namibia are characterised by relatively low concentrations and narrow ranges for CEC, clay and OC, which restricts the achievable explained variance and amplifies the effect of analytical uncertainty. In such contexts,

RMSE is a more informative metric for assessing model performance, as small absolute prediction errors can occur even when R^2 values are low.

5.4.9 Limitations and uncertainty

Despite the advances represented by the DSM products, limitations and sources of uncertainty must be recognised when interpreting the outputs. These challenges arise from the legacy nature of the observational dataset, limitations in covariates, and constraints associated with modelling approaches.

Spatial coverage of the observational database is uneven. Legacy data were collected over decades using heterogeneous sampling strategies and depths. Temporal mismatches between soil sampling dates and remote sensing imagery add uncertainty, particularly for dynamic properties (e.g. OC, N) that are sensitive to land management and land use changes, though this effect is expected to be minor under Namibia's low-intensity land use. The majority of analyses were performed by the ministerial Agriculture Laboratory using standardised methods, but the remaining analytical data were generated by other laboratories, introducing measurement uncertainty from differences in instrumentation, analytical procedures, and laboratory conditions. Depth-harmonisation, though necessary to provide internal consistency for DSM, added uncertainty, especially for at depth, and may underrepresent sharp horizon boundaries. Soil horizons rarely conform neatly to standardised depth intervals. In addition, some analytical variables exhibited skewed distributions with extreme outliers, reducing model robustness for those properties. These constraints are characteristic of many national-scale legacy datasets and are not unique to Namibia.

Downscaling and resampling of environmental covariates inevitably introduce uncertainty. Differences in spatial resolution, acquisition period, classification methodology, and preprocessing of covariates may propagate into the final soil predictions. Local categorical datasets (geology, soil types, vegetation, agro-ecological zones, carrying capacity) do not fully capture sub-90 m variability. Geological maps imperfectly represent regolith composition due to transport and mixing of parent materials. Climate data are gradually improving through the integration of remote sensing with ongoing weather station observations, though progress is hampered by historically sparse and poorly distributed station coverage, short recording periods for some attributes, and the changing climate. Global landcover and precipitation layers are themselves modelled products, with embedded spatial and temporal uncertainty. Terrain attributes and vegetation indices propagate any artefacts existing in the DEM and satellite imagery.

Machine learning offers powerful predictive tools, but present interpretation challenges as the mechanistic basis of predictions is not transparent. This study prioritised an operational, reproducible national DSM methodology and did not attempt comparison of different modelling algorithms. Random Forests is a proven, robust algorithm used in DSM and was selected for its stability and interpretability, although its predictive accuracy is lower than more sophisticated ensemble/hybrid approaches. For the modelling and uncertainty prediction, R 's wider range of ML algorithms and finer control over parameterisation were traded for Google Earth Engine's greater processing power and direct access to numerous global environmental datasets in the Earth Engine Data Catalogue. As GEE imposes user quotas on processing time, memory, and storage, a balance had to be struck regarding spatial resolution, model complexity, and the number of bootstrap iterations.

The uncertainty maps generated in this study constitute an important component of the DSM

framework rather than merely supplementary outputs. They highlight areas where predictions are less reliable and where additional sampling would most effectively improve future model performance. All PICP90 values approached or exceeded the ideal value, 90. These slight overestimations of uncertainty for some properties are preferable to overly narrow intervals. Users can interpret PI90 maps as conservative envelopes around the predicted means.

5.4.10 Complementary roles of conventional and digital soil mapping

Conventional and digital soil mapping address complementary aspects of soil information: CSM synthesises expert knowledge and heterogeneous legacy information into discrete soil map units, whereas DSM employs numerical methods to predict the continuous variation of individual soil properties within and across those units. The National Soil Map of Namibia delineates major soil-forming environments at national and regional scales and provides a practical stratification as input into DSM. It also serves as the benchmark against which DSM outputs are evaluated for plausibility and regional consistency, and it helps to identify areas where model outputs may be degraded by sparse data or covariate limitations.

5.5 The Namibian Soil Information System (NamSIS)

5.5.1 System structure

NamSIS integrates harmonised legacy profile and laboratory data, the national conventional soil map, and digital soil mapping outputs generated in this study within one conceptual framework (Figure 44). Original legacy materials are preserved alongside quality-controlled, format-standardised observational data and two sets of depth-harmonised data. Spatial data inputs and outputs are organised within a GIS environment, while map products are additionally provided in static image formats. Modelling scripts, analytical protocols, and metadata are included for transparency, reproducibility, and long-term stewardship.

The *GEOGRAPHICAL INFORMATION SYSTEM* and *DIGITAL SOIL MODELS* components spatially link the profile database with conventional soil mapping and predictive soil modelling products, for visualisation, spatial analysis, and reproducible modelling operations.

5.5.2 Geographical information system

The GIS includes vector (SHAPE files) and raster (GeoTIFF files) datasets in project workspaces, using QGIS open-source software. It supports data capture, storage, management, visualisation, mapping, spatial queries, and analysis.

The GIS environment includes:

- QGIS project files with preconfigured symbology and layouts;
- Vector layers representing:
 - NSPD observation locations;
 - NSMN soil map polygons and attributes;
 - Soil classes (WRB reference soil groups) and soil property layers derived from NSMN;

- Environmental covariates (polygon format);
- Administrative features such as boundaries, roads, and settlements;
- Raster layers including:
 - SRTM DEM and derived terrain attributes (e.g. slope, aspect, terrain position index, terrain ruggedness index);
 - DSM soil property prediction rasters (mean, 5th and 95th percentiles, and PI90);
 - Environmental covariate layers (grid format);
- Style files (QML) defining layer symbology and labels;
- Links to web services (WMS/WCS), including SoilGrids, Google and Bing satellite imagery, and OpenStreetMap basemaps, configured for on-demand access to external basemaps.

The geospatial infrastructure enables spatial analysis and the generation of new information products.

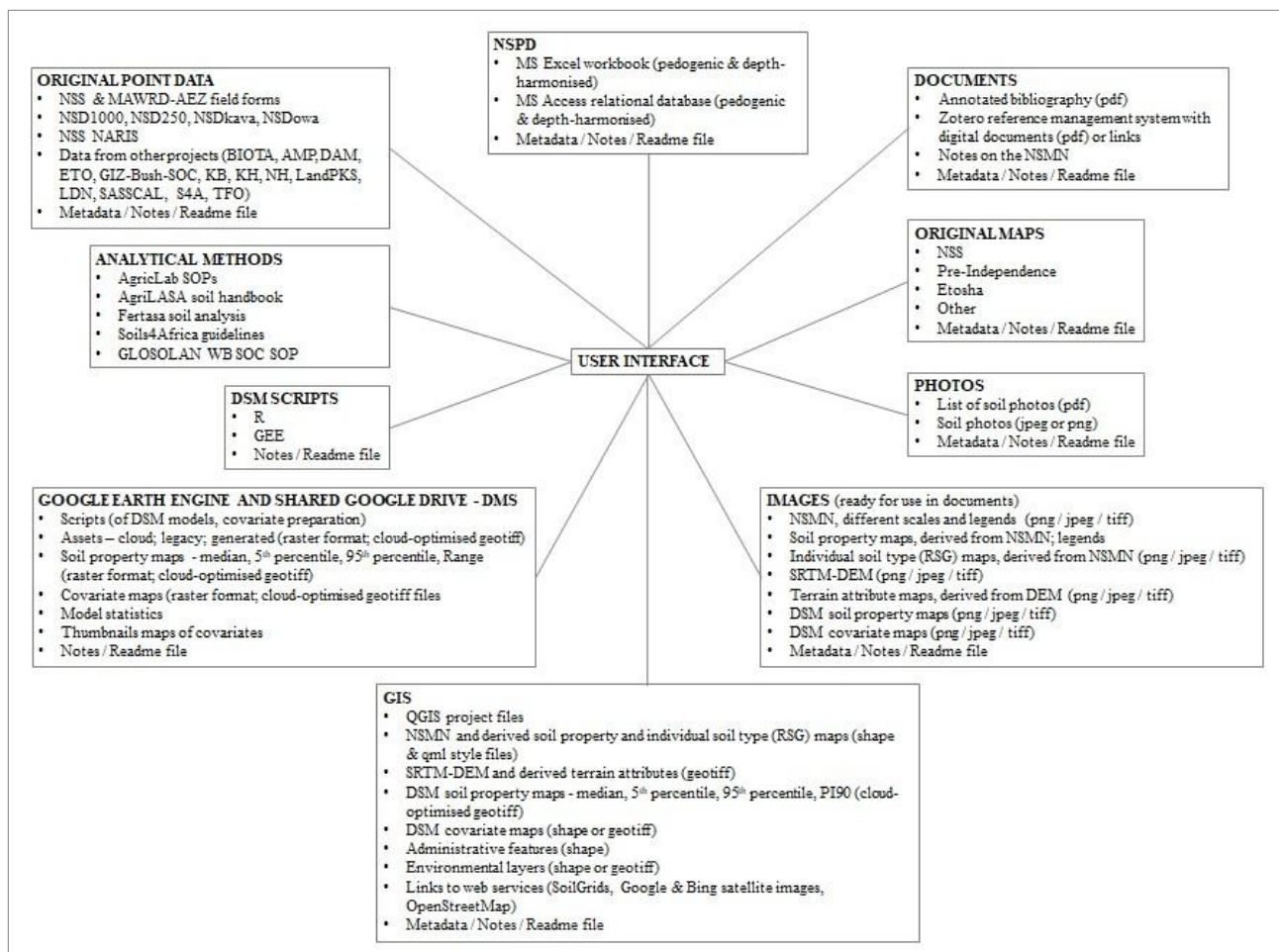


Figure 44. Components of the Namibian Soil Information System (NamSIS).

5.5.3 Digital soil modelling scripts

Digital soil modelling components are included in NamSIS as an organised collection of inputs, scripts, and outputs, as shown in the DSM Pipeline in Appendix D and available from (Gelsleichter, 2026a, 2026b; Gelsleichter & Coetzee, 2026). This structure ensures reproducibility and enables future model updates.

Inputs include depth-harmonised soil observations exported from the NSPD, environmental covariates, and terrain datasets prepared within Google Earth Engine (GEE).

The DSM operation is implemented using JavaScript, for execution on the Google Earth Engine (GEE) platform, and R. The module includes the R code for preprocessing and depth-harmonisation of soil data, hyperparameter tuning, feature selection, post-modelling metrics, and external validation. It also contains the JavaScript code for covariate preparation, model training and validation, and uncertainty quantification in GEE. Scripts are modular and semi-automated, allowing users with appropriate expertise to reproduce analyses, update predictions, or extend modelling to additional soil properties or depth intervals. Current models are configured for 90 m spatial resolution and three depth classes (0–30 cm, 30–60 cm, and 60–100 cm) but can be adapted to alternative depth schemes and resolutions. Aside from the static scripts, a link to a live model is provided for users to reproduce the modelling and recreate outputs using their own GEE accounts.

Final model outputs consist of cloud-optimised GeoTIFF raster layers representing prediction means, uncertainty bounds (5th and 95th percentiles), and 90% prediction intervals.

5.5.4 Image products

To support users without access to GIS software or modelling environments, key spatial products are provided as static image outputs (in PNG, JPEG, or TIFF formats) that can be inserted directly into reports and presentations.

Available image products include:

- NSMN at multiple typological levels;
- Maps of individual soil properties and reference soil groups derived from NSMN;
- Terrain models and derived indices;
- DSM predictions maps and associated uncertainty layers;
- Environmental covariate maps supporting conventional and digital mapping.

5.5.5 Metadata

Metadata documentation was compiled for the NSPD, NSMN, and DSM products following Namibia Spatial Data Infrastructure principles and relevant international standards. Metadata records include data provenance, attribute definitions, and spatial reference systems. This documentation supports future integration with regional and global soil information systems.

5.5.6 System design strengths and limitations

The modular system architecture allows ongoing integration and expansion as new datasets, analytical methods, and mapping products become available.

The inclusion of metadata and modelling scripts improves transparency and reproducibility, enabling future reanalysis or model refinement.

The integrated nature of NamSIS supports multi-purpose use of observations beyond their original survey purposes. Linked site and soil photographs provide visual context for morphological

interpretation, soil classification, and quality verification.

The adoption of consistent attribute definitions, harmonised data structures, and standardised metadata aligns NamSIS with FAIR principles and enables data exchange and integration with regional and global soil information systems. This increases the scientific value of Namibian soil data by allowing wider participation in global modelling efforts and comparative analyses.

Accessibility is increased through the provision of static map products, a variety of data formats (including PDF documents; PNG, JPEG and TIFF images; relational database, CSV and MS Excel spreadsheet tables; and SHAPE and GeoTIFF spatial files), and the use of open-source GIS software. Images offer fast, clear visualisation and communication for non-specialists, for example in the context of reporting, planning, and education. These deliberate design choices will increase wider stakeholder uptake and practical application of soil information.

Spatial distribution of soil data is uneven and particularly sparse in national parks, remote communal regions, and deeper subsoil layers. Some original field forms and datasets could not be recovered, preventing verification of those digital records and likely resulting in the permanent loss of some observational detail. Gaps in historical soil information cannot always be resolved. The inclusion of analytical method documentation reduces some of the challenges posed by the variation in analytical procedures and instrumentation among laboratories. This can help to identify and account for differences when interpreting trends or comparing values across datasets. However, perfect harmonisation is not always achievable. Uncertainty arising from these factors propagates through both conventional and digital soil mapping components. In DSM, it affects model calibration and prediction accuracy, while in CSM it influences the delineation and interpretation of soil units. Documenting these uncertainties is important for responsible use of the data.

5.5.7 Practical considerations and data access

Some design decisions required trade-offs between technical, institutional, and user-related considerations. Retaining raw as well as format-standardised and depth-harmonised data preserves provenance and supports future reanalysis but increases system complexity and data storage requirements.

Providing both relational database and spreadsheet formats of the profile database improves accessibility for diverse users but necessitates careful management to maintain synchronised versions. The system is to be maintained in two complementary versions: a public-access version intended for dissemination and general use, and a master version for safeguarding data integrity and managing dataset updates. This reflects a balance between open dissemination and data integrity and requires ongoing oversight.

MS Access was convenient for development of the database, but it is proprietary software and will pose problems for long-term user access, scalability, and sustainability. Migration to an open-source database management system will eventually be necessary, for example to PostgreSQL with the PostGIS extension that supports advanced spatial applications.

Long-term sustainability of a national soil information system requires stable institutional stewardship, technical capacity, and sustained commitment. National spatial and environmental information systems are most appropriately maintained within government institutions. In Namibia, responsibility for the National Spatial Data Infrastructure resides with the Namibia

Statistics Agency. However, institutional challenges relating to data accessibility, subject-specific curation, and dissemination infrastructure currently limit its suitability as a home for NamSIS. A university-based institutional model may provide a more viable medium-term option. Both the University of Namibia (UNAM) and the Namibia University of Science and Technology (NUST), through their expertise in geospatial sciences, natural resource management, agriculture, and environmental research, offer potentially suitable environments for hosting and further developing the system. Academic institutions are positioned well to support research-driven data integration, innovation, student training, and scientific collaboration. Nevertheless, university-based systems also face challenges related to staff turnover, shifting research priorities, short project cycles, and the potential loss of institutional memory over time. Until a permanent, institutional home for NamSIS is identified, a user interface on the World Wide Web (<https://soilsofnamibia.com/index.html>) provides links to the datasets and archives. This ensures immediate public accessibility and citation through DOI-based records. References of individual components, with associated DOI links, can be found in Appendix U .

While all components have been implemented and populated with available data, the creation of interactive web or mobile interfaces falls beyond the scope of this study.

6. CONCLUSIONS AND RECOMMENDATIONS

The study transformed Namibia's dispersed legacy soil information into a coherent national soil knowledge system that supports analysis, modelling, and long-term data stewardship.

6.1 Achieving the study's objectives

6.1.1 Development of a national soil information framework

The first objective was achieved through the development of the Namibian Soil Information System (NamSIS), a modular national framework linking quality-controlled soil observations, supporting archives, conventional and digital maps, metadata, and modelling resources. NamSIS provides information in formats suitable for specialist and non-specialist users and is designed to accommodate new observations, analytical methods, map revisions, and modelling outputs. It therefore establishes the scientific and structural basis for a cumulative national soil knowledge system rather than a collection of independent, project-specific products.

6.1.2 Legacy data recovery and consolidation

The second objective was achieved by recovering, digitising, cataloguing, and preserving Namibian soil documents, maps, datasets, field records, photographs, and laboratory method information that had previously been dispersed or at risk of loss. These resources were consolidated in structured digital archives, while 5 099 georeferenced observation sites and 13 425 soil horizons or layers were incorporated into the Namibian Soil Profile Database. Original records were retained alongside quality-controlled, standardised, and depth-harmonised versions, thereby preserving provenance while enabling verification, future reanalysis, and integration with contemporary mapping and modelling systems.

6.1.3 Conventional soil mapping

The third objective was achieved through the production of the National Soil Map of Namibia at a nominal scale of 1:250 000. The map provides complete national coverage, internal consistency, increased spatial and thematic detail, and alignment with the 2015 World Reference Base classification system. The map integrates heterogeneous legacy maps, soil profile observations, terrain information, environmental datasets, and expert interpretation within a coherent pedological framework. Derived thematic maps of key soil attributes further enhance the practical value of the dataset.

6.1.4 Digital soil modelling

The fourth objective was achieved by producing Namibia's first national machine-learning predictions of 16 key soil properties at 90 m resolution across three depth intervals. The performance of the Random Forests models trained on a depth-harmonised soil profile dataset was moderate to good for most properties, though prediction skill generally declined with depth. Predictions for sodium and electrical conductivity were weak. Spatial predictions were supported

by spatially explicit uncertainty estimates and qualitative pedological evaluation. Validation against an independent dataset showed that the national models performed comparable to or better than two global/continental soil products, highlighting the value of including a large amount of legacy soil data and locally produced environmental datasets.

Collectively, NamSIS, the Namibian Soil Profile Database, the National Soil Map, and the digital soil property predictions form Namibia's first comprehensive national soil information infrastructure. They provide a baseline for soil research, agricultural and rangeland management, ecosystem protection, land-degradation assessment, climate-change adaptation, environmental monitoring, and evidence-based natural resource governance.

6.2 Conceptual, methodological, and scientific contributions

Conceptually, the study demonstrates how soil knowledge can be integrated across complementary epistemological domains. By linking observational data, expert interpretation, and data-driven inference within a unified system, it provides a framework for transforming isolated survey outputs into a coherent and evolving national soil knowledge system capable of supporting both scientific investigation and practical land management.

Methodologically, the study developed a transferable framework for constructing national soil information systems in data-limited countries. The modular DSM workflow provides a reproducible and extendable methodology for iterative updates and facilitates future incorporation of new observations, covariates, and modelling techniques.

Scientifically, the research contributes to pedological understanding by confirming that soil variability across Namibia is strongly structured by climatic gradients, topography, and parent material. These controls are expressed through characteristic dryland pedogenic processes. The consistency between conventional soil mapping patterns and DSM outputs provides empirical support for the applicability of SCORPAN-based modelling in arid and semi-arid environments. The study supports the value of including national legacy data and expert-derived, context-specific environmental predictors that capture local variance and conditions in DSM.

The study engages with current debates and emerging directions in soil information management and mapping. It supports the growing consensus that hybrid approaches combining conventional soil mapping and digital modelling produce more robust and interpretable results than either approach alone. By addressing persistent challenges related to data scarcity, inconsistency, and limited interoperability, and by embedding soil observations, maps, models, and metadata within a FAIR-aligned system, the research aligns with contemporary priorities in transparency, reproducibility, and system integration. The results reinforce the importance of context-specific, bottom-up approaches to soil mapping. The inclusion of uncertainty estimates, modular workflows, and extensible system design positions the work within current efforts to develop dynamic, updateable, and transferable soil information systems.

6.3 Hypothesis support

6.3.1 Hypothesis H₁ — Data integration

The consolidation of a large amount of previously scattered legacy soil information into a central

repository, the transformation of analogue materials into widely used digital formats, and the use of open-source GIS software collectively improve accessibility for a wide range of users. Standardisation of data structures and depth harmonisation enhance compatibility with international soil information products and modelling frameworks. The inclusion of metadata, modelling scripts, and laboratory protocols improves transparency and reproducibility, enabling users to understand and replicate the processes through which the data and derived products were generated. The geospatial infrastructure — the GIS and DSM components — expands the analytical potential of the system by enabling spatial analysis and the generation of new information products. The archival preservation of grey literature, photographs, field notes, original datasets, and legacy maps safeguards historical information alongside standardised, harmonised and enhanced datasets. Collectively, the integration of these components significantly improves the accessibility, standardisation, long-term preservation, and practical utility of Namibia's soil information — supporting Hypothesis H₁.

6.3.2 Hypothesis H₂ — Pedogenic-environmental relationships

Hypothesis H₂ relates to the SCORPAN framework that underpins DSM. It is supported by the modelling results obtained in this study. The environmental covariates that emerged as the most influential predictors — particularly those related to climate, terrain, vegetation and proxies of parent material — correspond closely with the main soil-forming factors governing pedogenesis in dryland environments. The spatial patterns of both predicted soil properties and those emerging from the conventional map align with major climatic gradients, physiographic regions, and known pedogenic processes across the country. More than half of the DSM models incorporated locally derived covariates, demonstrating that local datasets contain process-relevant information that are not necessarily captured by global proxies. These results indicate that environmental factors exert strong controls on soil variability at national scale and confirm that such relationships can be applied to predict soil properties where direct observations are sparse.

6.3.3 Hypothesis H₃ — Conventional-digital soil mapping convergence

The refinement of the conventional soil map through the integration of legacy observations and maps, expert interpretation, and ancillary spatial datasets improved the delineation of soil boundaries and increased both the spatial and typological detail of the national map. The updated map now covers the entire country and is more internally consistent than any previous national products. At the same time, the conventional soil map and other locally produced environmental maps enhanced DSM performance as meaningful categorical predictors and by offering a pedological framework for evaluating predicted spatial patterns. In turn, DSM contributes quantitative information on soil properties and reveals subtle environmental gradients that can guide iterative improvements to the conventional soil map. Together, these results demonstrate the advantages of a hybrid mapping strategy in which expert knowledge and machine-learning approaches reinforce one another, producing more internally coherent and interpretable soil information than either method alone, particularly in data-sparse dryland environments — confirming Hypothesis H₃.

6.3.4 Overarching hypothesis

The three subsidiary hypotheses are supported through the successful development of NamSIS (H_1), the predictive performance of the DSM models (H_2), and the demonstrated complementarity between conventional and digital soil mapping approaches (H_3). In addition, the external validation results are consistent with improved predictive performance compared with the global and continental products that were previously the only available spatial coverages of soil properties for Namibia, thereby directly supporting the overarching hypothesis:

Integrating Namibia's fragmented legacy soil information with global and locally derived environmental covariates — through harmonisation, conventional soil mapping, and machine-learning-based digital soil mapping — produces more consistent, accessible, and nationally useful soil information than previously available.

6.4 Recommendations for future work

The work presented in this dissertation represents a major step forward in soil information management in Namibia, but should not be regarded as a final, static product. It is the departure point for developing operational, interactive web- and mobile-based applications, and for continued refinement and expansion as new data, methods, and technologies become available.

6.4.1 Improving the observational basis

Although the Namibian Soil Profile Database now contains more than twice the number of georeferenced soil profiles/augerings described during the 1998–2000 National Soil Survey, large parts of Namibia remain under-sampled. Future mapping and sampling campaigns should be strategic rather than opportunistic. They should be guided by analyses of current observation density and distribution, DSM uncertainty patterns, and the representation within soil classes, agro-ecological zones, and other relevant environmental strata. These efforts should ideally follow purposive stratified random designs (Brus, 2019; Brus et al., 2011). The homosoil concept (Mallavan et al., 2010; Nenkam et al., 2023; Pásztor et al., 2014) and pseudo-observations (Hengl et al., 2017b) could augment sparse datasets. New surveys should be planned from the outset for integration into NamSIS, using standardised field descriptions, laboratory protocols, and metadata capture. Filling gaps in soil data will benefit both conventional and digital soil mapping.

6.4.2 Enhancing the national soil map

Future refinement can be achieved by targeted field verification in areas of high uncertainty or high land-use importance and formal accuracy assessment using independent datasets.

Spatial delineation and internal composition of soil map units can be progressively refined as new soil observations and improved geological maps become available. The legend structure, STUs, and SMUs should be reviewed periodically and kept aligned with updates to the WRB classification system.

Clear accompanying documentation and interpretive information will help non-specialist users interpret the map more effectively, increasing the practical value of the map. Map-unit-based

decision-support tools (e.g. for crop suitability, land capability classification, or erosion risk assessment) can provide a more intuitive understanding of soil capabilities and limitations.

The NSMN represents a snapshot in time. Environmental change, land-use change (e.g. mining, urban expansion), land degradation (e.g. deforestation, erosion, salinisation), and human interventions (e.g. rangeland rehabilitation) will alter soil conditions over time. Establishing a national soil monitoring network would enable long-term evaluation of soil trends and assessment of management and climate impacts. Citizen science and local knowledge from farmers, extension officers, and park rangers could assist in identifying areas with unusual soil conditions or rapid change, flagging soil map units that may need periodic updating.

6.4.3 Advancing digital soil mapping

The DSM framework benefited from combining generic global datasets with locally derived covariates, and further advances are possible. Enhanced climate products can be incorporated as station networks expand and improve, and new satellite-derived products are developed. Radar-based remote sensing could be added to better capture surface roughness, moisture, and structure. Mineralogical indices derived from airborne or satellite hyperspectral data should be included as new datasets are released. Further research is needed to develop maps of parent material, regolith thickness, and degree of weathering. Emerging geospatial feature-embedding datasets (Brown et al., 2025) also warrant investigation.

This study prioritised an operational procedure using Random Forests in Google Earth Engine rather than algorithm comparison. Future research could benchmark RF against alternative approaches such as gradient boosting, quantile regression forests, stacked ensembles, or deep learning methods. Future studies could assess alternative uncertainty quantification strategies and investigate the propagation of measurement and depth-harmonisation errors.

The current DSM outputs are effectively anchored to the temporal context of the covariates and legacy data. Future work could explore time-series DSM using multi-date remote sensing and repeated soil measurements. This would enable analysis of drought and wet-year impacts and land-use change and allow modelling of soil trajectories under different management and climate scenarios — moving DSM from static mapping towards dynamic monitoring and forecasting.

There is a need to translate digital soil mapping outputs into products for specific user groups. Web-based tools would allow lay users to query predicted soil properties and associated uncertainty at point or field scale. Training materials will be important for integrating DSM into extension services. The next major stage of soil information development in Namibia would be to move beyond mapping to decision-support applications, such as nutrient management planning, and assessing and monitoring ecosystem functioning. Tailored products are needed to support policy frameworks such as land degradation neutrality and climate-smart agriculture.

6.4.4 Further developing NamSIS and user access

The modular design of NamSIS allows expansion to accommodate emerging forms of soil information. These could include: proximal and field spectroscopy (e.g. vis–NIR and MIR) for rapid and cost-effective soil characterisation at additional locations and depths; geophysical data (e.g. electromagnetic conductivity, gamma radiometrics, airborne hyperspectral data, ground-

penetrating radar); biological and microbiological indicators (e.g. soil microbial biomass, environmental DNA, and diversity indices); and repeated measurements at permanent monitoring plots to capture temporal dynamics (e.g. SOC change, salinisation, erosion, rehabilitation outcomes). Clear protocols will be needed to link these data to existing profiles and map units, so they become fully integrated components rather than parallel datasets.

Future development should promote transparency and usability by continual updating and expanding metadata across all system components in line with evolving national and international standards. Data-quality flags, version control mechanisms, and clear distinctions between raw observations, harmonised data, and modelled outputs should be maintained.

User-friendly interfaces, such as mobile applications, online map viewers, interactive WebGIS applications, and search, query and download tools, together with supporting documentation (e.g. data dictionaries and sector-specific user guides) will ensure that NamSIS functions not only as a scientific repository but also as a practical tool for non-specialist users. Future work could include automated tools for data ingestion and quality control, formatting into other machine-readable formats (e.g. JSON, XML), creation of versioning and transformation logs, and access authorisation tools.

6.4.5 Institutional issues

The long-term sustainability of NamSIS requires incorporation into the National Spatial Data Infrastructure. Collaborative institutional arrangements must be supported by clear mandates and responsibilities, governance structures, professional curation, and ongoing investment in soil information management and capacity development.

Soil scientists, GIS specialists, and remote sensing analysts need training in DSM methods and data management, while capacity must be built among extension officers, planners, and environmental managers to interpret and apply soil information. Collaboration between universities, ministries, and research institutes is required to sustain expertise across generations.

The legacy nature of the current database means that differences between laboratories and analytical methods remain a source of uncertainty. Continued support for laboratories is essential to maintain standardised methodologies, strengthen internal quality assurance, and participate in proficiency testing schemes.

The link between soil information and policy processes must be strengthened. This includes embedding NamSIS outputs within national frameworks addressing, for example, rural development, land degradation neutrality, climate change adaptation, agricultural planning, and biodiversity conservation. Indicators and products should be co-developed with stakeholders to ensure that soil information responds directly to practical policy needs.

Zenodo and a website provide a workable short- to medium-term solution for public dissemination, DOI-based citation, and preservation of static releases of NamSIS datasets. Future development of NamSIS should include migration from archival storage toward a dynamic server-based architecture using systems such as PostgreSQL/PostGIS for spatial database management and GeoServer or QGIS Server for web-based geospatial data dissemination and services.

The preservation of physical soil samples is of great concern. Samples collected during the NSS and MAWF–AEZ campaigns are uncatalogued, unsecured, and some have already been discarded. Historical soil samples are a valuable national asset and must be securely preserved. A formal

archiving policy and protocol — defining the responsibilities, acquisition-, storage- and disposal modalities, retention period, cataloguing, and sharing of soil samples, etc. — is urgently needed for long-term custodianship.

Namibia forms part of a broader regional and global soil information community. The work presented here provides a foundation for continued contributions to global and continental initiatives such as WoSIS, SoilGrids, the Africa Soil Information System, and the African Union Soil Observatory, as well as participation in regional efforts addressing dryland soils, carbon accounting, land degradation, and food security. Sharing methods, scripts, and experiences with other data-scarce countries will contribute to collective progress.

6.5 Final remarks

The work illustrates that meaningful advances in soil science depend not only on new data or models, but on the effective integration and stewardship of existing knowledge. By establishing such a unified, multi-layered soil information architecture, the study contributes toward a coherent, sustainable foundation for soil knowledge in Namibia, and for past observations to remain available to inform future understanding. It also provides a transferable framework for developing soil information systems in data-limited regions, supporting both scientific advancement and evidence-based land management.

7. NEW SCIENTIFIC RESULTS

1. This research **developed the first integrated national soil knowledge infrastructure for Namibia** which brings quality-controlled soil observations, legacy data archives, conventional soil maps, digital soil maps with uncertainty assessments, metadata, photographs, analytical protocols, and reproducible modelling scripts into a single FAIR-aligned, extensible framework.
2. The study **rescued and preserved Namibia's at-risk soil science legacy** by recovering, verifying, digitising, cataloguing, and publicly archiving historical soil datasets, maps, field records, photographs, laboratory methods, reports, and grey literature that had previously been dispersed and largely inaccessible.
3. The research produced the **first complete and internally coherent National Soil Map of Namibia at a scale of 1:250 000**, providing full territorial coverage — including previously unmapped areas — and substantially increasing spatial and typological detail to 4 787 polygons, 212 Soil Typological Units, and 490 Soil Map Units classified according to the 2015 World Reference Base for Soil Resources.
4. The study produced the **first comprehensive, quality-controlled, standardised, publicly accessible Namibian Soil Profile Database**, consolidating 5 099 georeferenced observation sites and 13 425 soil horizons or layers, in addition to depth-harmonised versions suitable for modelling.
5. The research generated **Namibia's first spatially continuous national predictions of physical and chemical soil properties at 90 m resolution** for the 0–30, 30–60, and 60–100 cm depth intervals, together with prediction means, 5th and 95th percentile uncertainty boundaries, and 90% prediction intervals. Independently validation produced comparable or better metrics than global and pan-African datasets. It was shown that locally developed environmental datasets contain substantial process-relevant information and should be used to supplement generic global covariates in national digital soil mapping.
6. The study **validated a hybrid conventional–digital soil-mapping strategy** in which expert-derived conventional maps provide environmental knowledge, spatial stratification, and pedological plausibility checks for machine-learning models, while digital predictions quantify continuous within-unit variation and identify priorities for conventional map refinement.
7. The study **established a foundation for soil research and practical applications in Namibia** that did not exist before — a baseline for agricultural and rangeland management, ecosystem protection, land-degradation assessment, climate-change adaptation, environmental monitoring, and evidence-based natural resource governance.
8. The research **developed a methodology for constructing national soil information systems that is reproducible and transferable to other countries**. The methodology integrates legacy data revitalisation, database harmonisation, conventional mapping, machine-learning spatial prediction, uncertainty quantification, and external validation within a comprehensive, interconnected framework.

SUMMARY

Reliable soil information is essential for agricultural development, sustainable land management, climate-change adaptation, biodiversity conservation, environmental assessment, land-degradation monitoring, and natural-resource governance. In Namibia, however, soil information has historically been fragmented among government institutions, research organisations, development projects. Much of this information was difficult to locate, incompletely documented, stored in obsolete or vulnerable formats, and unsuitable for direct integration into contemporary geographical information systems and digital soil models. The absence of a complete national soil map, a harmonised national soil profile database, and spatially explicit soil property maps has constrained the use of soil information in research, planning, policy formulation, and environmental reporting.

This study addressed these deficiencies by transforming the country's dispersed legacy soil information into an integrated national soil knowledge infrastructure. Four interrelated objectives were pursued: the conceptualisation and development of the Namibian Soil Information System (NamSIS); the identification, rescue, digitisation, and consolidation of legacy soil information; the production of an updated and enhanced national conventional soil map classified according to the 2015 World Reference Base for Soil Resources; and the development of spatially continuous predictions of selected soil properties using machine-learning methods.

The workflow started with data archaeology, rescue and renewal. Historical soil datasets, maps, reports, field records, photographs, laboratory records, and grey literature were recovered from government institutions, universities, research organisations, consultants, and development projects. Field records, legacy maps, photographs, and documents were assessed, digitised, catalogued, quality controlled, and incorporated into structured, publicly accessible archives. Original source materials were retained to safeguard Namibia's scientific memory, allow future users to verify data provenance and processing decisions, and create opportunities for reinterpretation as analytical, classification, and modelling methods develop.

Profile descriptions, geographical coordinates, diagnostic properties, sampling depths, analytical methods, units of measurement, and laboratory results were verified and standardised. The resulting Namibian Soil Profile Database contains 5 099 georeferenced observation sites and 13 425 soil horizons or layers, more than doubling the number of nationally available georeferenced observations compared with the 1998–2000 National Soil Survey. The database integrates observations from 34 surveys and projects undertaken over several decades. It retains original records alongside quality-controlled and format-standardised data and two depth-harmonised datasets that were prepared for spatial analysis and modelling. Associated dataset documentation, laboratory protocols, photographs, and metadata improve the transparency, traceability, and reuse potential of the observations. Due to the legacy nature of the data, the spatial distribution is uneven: observations are concentrated in accessible and agriculturally important regions and are comparatively sparse in national parks, remote communal areas, rugged terrain, desert environments, and deeper soil layers.

Conventional soil mapping combined legacy maps, profile observations, geological and terrain information, satellite imagery, environmental datasets, and expert pedological interpretation to produce the new National Soil Map of Namibia. The map provides, for the first time, complete national coverage, including areas that were previously unmapped or represented only at very small scales, such as the Skeleton Coast, Tsau Ikhueb, and Etosha National Parks. It is aligned

with the 2015 edition of the World Reference Base for Soil Resources and comprises 490 Soil Map Units composed of 212 Soil Typological Units. Map units represent dominant, associated, and included soils, thereby recording within-unit complexity realistically. The new National Soil Map of Namibia represents a substantial improvement over the Preliminary Soil Map produced through the 1998–2000 National Soil Survey. It provides greater spatial and typological detail, fuller representation of soil qualifiers, and stronger links to soil-profile and environmental data. Its attribute structure includes information relevant to land management. Derived maps of individual Reference Soil Groups, soil characteristics and terrain attributes increase its interpretive and practical value. The map is intended principally for national and regional applications and can be used with high confidence at scales smaller than 1:250 000 and moderate confidence at scales between that and 1:100 000. Apparent detail at larger display scales should not be interpreted as greater positional or thematic accuracy, particularly in areas supported by sparse observations or reconnaissance-level source information.

The conventional mapping results show strong correspondence between soil distribution and Namibia's major climatic, geomorphological, and lithological regions. Deep, quartz-rich Arenosols dominate extensive parts of the Kalahari Basin, whereas Leptosols, Regosols, Calcisols, Cambisols, and other soil groups form more complex mosaics across the Escarpment, Central Highlands, Kaokoveld, Damaraland, and southern Namibia. Limited leaching, high evapotranspiration, carbonate and salt accumulation, aeolian sand deposition, erosion, and sediment redistribution are major controls on soil development. The map provides not only national coverage but also a coherent pedological interpretation of Namibia's dryland landscapes.

Digital soil mapping used a modelling subset of 4 958 records extracted from the Namibian Soil Profile Database. Heterogeneous horizon data were harmonised to three agriculturally relevant depth intervals: 0–30, 30–60, and 60–100 cm. Sixteen physical and chemical soil properties were modelled: bulk density, base saturation, pH measured in water, cation-exchange capacity, extractable calcium, magnesium, potassium and sodium, electrical conductivity determined by two methods, organic carbon, plant-available phosphorus, total nitrogen, sand, silt, and clay. A stack of 65 environmental covariates was assembled, representing climate, terrain, hydrology, geology, landforms, vegetation, land cover, remote-sensing bands and indices, agro-ecological zones, carrying capacity, and the newly developed national soil map. Covariate selection combined automated Boruta analysis with expert review, while Random Forest hyperparameters were optimised separately for each soil-property and depth combination. Twenty bootstrap iterations were used for every property and depth interval, resulting in 960 individual models. The models generated national predictions at 90 m resolution, and the mean of the bootstrap realisations was used as the central prediction for each property and depth.

Model performance varied considerably among properties and depths, with coefficients of determination ranging from 0.10 to 0.75. Moderate to strong predictive performance was achieved for properties with strong climatic, lithological, textural, or vegetation controls, including pH, base saturation, cation-exchange capacity at shallower depths, total nitrogen, sand, silt, and exchangeable calcium, magnesium and potassium. Predictive skill generally declined with depth because subsoil observations were less numerous and subsoil conditions were less directly related to surface-derived environmental covariates. Organic carbon and phosphorus were less predictable because they are affected by local biological cycling, land management, and differences in analytical methods. The weak performance of sodium and electrical conductivity models can be attributed to strong skewness of the training data from agriculture-oriented legacy surveys avoiding saline areas plus a small number of extreme values from salt pans.

The predicted spatial patterns were generally pedologically plausible. Many properties displayed a prominent transition between the deep, relatively homogeneous, quartz-rich sands of the Kalahari Basin and the more geologically diverse soils of the Escarpment, Central Highlands, and western, central, and southern regions. Higher rainfall and coarse texture in the northeast promote leaching and comparatively lower pH, base saturation, cation-exchange capacity, and exchangeable-base concentrations. Drier areas underlain by calcareous or base-rich parent materials generally display higher pH and greater base accumulation. Organic carbon and nitrogen are comparatively elevated in the more humid northeast and along ephemeral river systems, while finer textures occur in alluvial systems and landscapes underlain by basaltic, calcareous, dolomitic, schistose, and related materials.

Covariate selection and variable importance results confirmed that climate, elevation, vegetation, hydrology, parent material, and broad terrain structure exert strong controls on national soil variability. Locally produced soil, vegetation, geology, agro-ecological-zone, and carrying-capacity maps were included in more than half of the models and often ranked among the influential predictors. Their importance suggests that nationally developed environmental products contain process-relevant knowledge that is not fully represented by generic global covariates.

Prediction uncertainty was quantified spatially using the 5th and 95th percentiles of the bootstrap predictions and the resulting 90% prediction-interval widths. Uncertainty generally increased with depth and was greater in sparsely sampled or environmentally heterogeneous areas. Prediction Interval Coverage Probability values were close to or above the nominal 90% target for all properties, indicating that the intervals provide realistic or moderately conservative representations of uncertainty. The uncertainty maps are important operational products because they distinguish areas where predictions are comparatively reliable from those requiring cautious interpretation and provide an objective basis for targeting future field observations.

Independent external validation was undertaken using Namibian observations from the Soils4Africa project that were excluded from model calibration. Predictions from this study were compared with SoilGrids250m and iSDA products. Although the independent dataset was small and unevenly distributed, the Namibian models generally produced coefficients of determination comparable to or higher than those of the reference products and consistently achieved comparable or lower root mean squared errors. Improvements were particularly evident for particle-size fractions. These results demonstrate the predictive benefit of incorporating expanded national legacy observations and locally relevant environmental covariates rather than relying exclusively on continental or global products.

The conventional and digital mapping components proved complementary. The conventional soil map, embedding expert knowledge, supplied locally relevant predictors for digital modelling and provided a pedological framework for evaluating the plausibility of model outputs. Digital soil mapping supplied spatially continuous quantitative predictions and uncertainty assessments within and across conventional map units and identified areas in which additional observations could improve both modelling and conventional mapping. Their integration produced a more comprehensive, internally coherent, and interpretable representation of Namibia's soil resources than either approach could have provided independently.

NamSIS brings these outputs together within a modular national framework comprising the soil profile database, documentary and photographic archives, the soil map, digital soil property predictions with associated uncertainty measures, modelling scripts, environmental layers,

geographical information system projects, static image products, and metadata. Information is supplied in multiple formats to support specialist and non-specialist users. The system is aligned with FAIR principles and is designed to accommodate new observations, analytical methods, map revisions, and modelling outputs. However, its long-term sustainability will depend on stable institutional stewardship, professional data curation, technical capacity, and migration to an open, server-based spatial database and web-service architecture.

The study supports the operational hypotheses. The development of NamSIS demonstrates that harmonising and integrating fragmented legacy information substantially improves its accessibility, interoperability, preservation, and utility. The digital modelling results confirm that national soil variability can be predicted from climate, terrain, vegetation, parent-material, and related environmental gradients. The reciprocal contributions of conventional and digital mapping confirm that expert interpretation and machine-learning inference are mutually reinforcing. Collectively, these findings support the overarching proposition that integrating Namibia's legacy soil information with global and locally derived environmental covariates produces more consistent, accessible, and nationally useful soil information than was previously available.

The principal scientific contributions are therefore the establishment of Namibia's first integrated national soil knowledge infrastructure; creation of its first comprehensive public soil profile database; rescue and preservation of an irreplaceable national soil-data legacy; production of the first complete and cohesive national soil map; generation of the first depth-specific national predictions and uncertainty maps for key soil properties; improved understanding of the environmental controls on Namibian dryland pedogenesis; empirical validation of a hybrid conventional–digital mapping strategy; development of a reproducible framework that can be adapted for soil-information development in other data-scarce countries; and establishment of a baseline for soil research and evidence-based decision-making in Namibia.

APPENDICES

A. References

- Abugri, B., Kombat, R., & Pennaneach, S. (2025). *African Union Soil Observatory project officially launched to close Africa's soil data gaps for sustainable agriculture*. African Union Soil Observatory. <https://auso.faraafrica.org/2025/10/06/african-union-soil-observatory-project-officially-launched-to-close-africas-soil-data-gaps-for-sustainable-agriculture/>
- Adeniyi, O. D., Bature, H., & Mearker, M. (2024). A systematic review on digital soil mapping approaches in lowland areas. *Land*, 13(3), 379. <https://doi.org/10.3390/land13030379>
- Adhikari, K., Hartemink, A.E., Minasny, B., Bou Kheir, R., Greve, M.B., & Greve, M.H. (2014). Digital mapping of soil organic carbon contents and stocks in Denmark. *PLOS ONE* 9, e105519. <https://doi.org/10.1371/journal.pone.0105519>
- Adhikari, K., & Hartemink, A. E. (2016). Linking soils to ecosystem services—A global review. *Geoderma*, 262, 101–111. <https://doi.org/10.1016/j.geoderma.2015.08.009>
- African Union Assembly. (2003). *Maputo Declaration on agriculture and food security in Africa*. Second Ordinary Assembly of the African Union.
- African Union Commission. (n.d.). African Union Soil Observatory (AUSO): Project overview. *Africa Union Soil Observatory*. Retrieved 14 May 2026, from <https://auso.faraafrica.org/overview-2/>
- AgriLASA. (2004). *AgriLASA soil handbook*. Agri Laboratory Association of Southern Africa.
- Ahrari, A. (2024). *ASTER lithology* [Computer software]. https://github.com/AmirhosseinAhrari/GoogleEarthEngine/blob/main/00067_aster_lithology
- Angelopoulou, T., Tziolas, N., Balafoutis, A., Zalidis, G., & Bochtis, D. (2019). Remote Sensing Techniques for Soil Organic Carbon Estimation: A Review. *Remote Sensing*, 11(6), 676. <https://doi.org/10.3390/rs11060676>
- AOC Technical Services (Pty) Ltd. (1966). *Report on the preliminary reconnaissance soil survey of the Kunene River Irrigation Project*.
- AOC Technical Services (Pty) Ltd. (1967a). *A preliminary survey of the natural environment and the agricultural resources of Okavangoland*.
- AOC Technical Services (Pty) Ltd. (1967b). *A proposal for an irrigation research scheme in the vicinity of the canal route in the region east of the oshana Olusandja—Owamboland*.
- AOC Technical Services (Pty) Ltd. (1967c). *Report on a reconnaissance soil survey of Okavangoland with particular reference to the riparian land*.
- Aroca-Fernandez, J. M., Diez-Pastor, J. F., Latorre-Carmona, P., Elvira, V., Camps-Valls, G., Pascual, R., & Garcia-Osorio, C. (2025). *A collaborative platform for soil organic carbon inference based on spatiotemporal remote sensing data* (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.2504.13962>
- Arrouays, D., Grundy, M. G., Hartemink, A. E., Hempel, J. W., Heuvelink, G. B. M., Hong, S. Y., Lagacherie, P., Lelyk, G., McBratney, A. B., McKenzie, N. J., Mendonca-Santos, M. d.L., Minasny, B., Montanarella, L., Odeh, I. O. A., Sanchez, P. A., Thompson, J. A., & Zhang, G.-L. (2014). GlobalSoilMap. In *Advances in Agronomy* (Vol. 125, pp. 93–134). Elsevier. <https://doi.org/10.1016/B978-0-12-800137-0.00003-0>
- Arrouays, D., Lagacherie, P., & Hartemink, A. E. (2017a). Digital soil mapping across the globe. *Geoderma Regional*, 9, 1–4. <https://doi.org/10.1016/j.geodrs.2017.03.002>
- Arrouays, D., Leenaars, J. G. B., Richer-de-Forges, A. C., Adhikari, K., Ballabio, C., Greve, M., Grundy, M., Guerrero, E., Hempel, J., Hengl, T., Heuvelink, G., Batjes, N., Carvalho, E., Hartemink, A., Hewitt, A., Hong, S.-Y., Krasilnikov, P., Lagacherie, P., Lelyk, G., ... Rodriguez, D. (2017b). Soil legacy data rescue via GlobalSoilMap and other international and national initiatives. *GeoResJ*, 14, 1–19. <https://doi.org/10.1016/j.grj.2017.06.001>
- Arrouays, D., McBratney, A., Bouma, J., Libohova, Z., Richer-de-Forges, A. C., Morgan, C. L., Roudier, P., Poggio, L., & Mulder, V. L. (2020a). Impressions of digital soil maps: The good, the not so good, and making them ever better. *Geoderma Regional*, 20, e00255. <https://doi.org/https://doi.org/10.1016/j.geodrs.2020.e00255>
- Arrouays, D., McKenzie, N., Hempel, J., Richer-de-Forges, A., & McBratney, A. (Eds). (2020b). *GlobalSoilMap: Digital soil mapping from country to globe*. CRC Press. <https://doi.org/https://doi.org/10.1201/9781351239707>
- Arrouays, D., Poggio, L., Salazar Guerrero, O. A., & Mulder, V. L. (2020c). Digital soil mapping and GlobalSoilMap. Main

- advances and ways forward. *Geoderma Regional*, 21, e00265. <https://doi.org/10.1016/j.geodrs.2020.e00265>
- Atlas of Namibia Team. (2022). *Atlas of Namibia: Its land, water and life*. Namibia Nature Foundation.
- Batjes, N. H. (2002). *Soil parameter estimates for the soil types of the world for use in global and regional modelling (Version 2.1)*. International Food Policy Research Institute (IFPRI), International Soil Reference and Information Centre (ISRIC). <https://edepot.wur.nl/3075>
- Batjes, N. H. (2004). *SOTER-based soil parameter estimates for southern Africa*. International Soil Reference and Information Centre (ISRIC). https://isric.org/sites/default/files/isric_report_2004_04.pdf
- Batjes, N. H. (2008). Quality assessment of soil profile data. *Geoderma*, 148, 1–12.
- Batjes, N. H., Calisto, L., & De Sousa, L. M. (2024). Providing quality-assessed and standardised soil data to support global mapping and modelling (WoSIS snapshot 2023). *Earth System Science Data*, 16(10), 4735–4765. <https://doi.org/10.5194/essd-16-4735-2024>
- Batjes, N. H., & Ribeiro, E. (2019). *Updated MS-Access SOTER template for the EU Danube basin project with worked examples (Deliverables A1, A2 and A4, EU JRC Service Contract Nr. 936108, Collection of soil data in SOTER format from 14 Danube strategy countries, at scale 1:250 000)* (Project Report No. 1). ISRIC–World Soil Information.
- Batjes, N. H., Ribeiro, E., & Van Oostrum, A. (2020). Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019). *Earth System Science Data*, 12(1), 299–320. <https://doi.org/10.5194/essd-12-299-2020>
- Batjes, N. H., Ribeiro, E., Van Oostrum, A., Leenaars, J., Hengl, T., & Mendes De Jesus, J. (2017). WoSIS: Providing standardised soil profile data for the world. *Earth System Science Data*, 9(1), 1–14. <https://doi.org/10.5194/essd-9-1-2017>
- Baxter, S. J., & Crawford, D. M. (2008). Incorporating legacy soil pH databases into digital soil maps. In A. E. Hartemink, A. McBratney, & M. D. L. Mendonça-Santos (Eds), *Digital Soil Mapping with Limited Data* (pp. 311–318). Springer Netherlands. https://doi.org/https://doi.org/10.1007/978-1-4020-8592-5_27
- Bazaglia Filho, O., Rizzo, R., Lepsch, I. F., Prado, H. D., Gomes, F. H., Mazza, J. A., & Demattê, J. A. M. (2013). Comparison between detailed digital and conventional soil maps of an area with complex geology. *Revista Brasileira de Ciência Do Solo*, 37(5), 1136–1148. <https://doi.org/10.1590/S0100-06832013000500003>
- Behrens, T., Schmidt, K., MacMillan, R. A., & Viscarra Rossel, R. A. (2018). Multi-scale digital soil mapping with deep learning. *Scientific Reports*, 8(1), 15244. <https://doi.org/10.1038/s41598-018-33516-6>
- Benin, S., & Yu, B. (2013). Complying with the Maputo Declaration target: Trends in public agricultural expenditures and implications for pursuit of optimal allocation of public agricultural spending. In *ReSAKSS Annual Trends and Outlook Report 2012*. International Food Policy Research Institute (IFPRI).
- Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimisation. *Journal of Machine Learning Research*, 13, 281–305.
- Beugler-Bell, H. (1991). *Untersuchungen zur bodenerosion im Etoscha Nationalpark, Namibia—Unter besonderer berücksichtigung der erodierbarkeit der böden*. [Diplomarbeit]. Regensburg University.
- Beugler-Bell, H., & Buch, M. W. (1997). Soils and soil erosion in the Etosha National Park, northern Namibia. *Madoqua*, 20(1), 91–104.
- Beugler-Bell, H., Buch, M. W., & Trippner, C. (1993a). *A guideline for soil classification in the Etosha National Park and adjacent areas in central northern Namibia based on the FAO/UNESCO legend of the soil map of the world and including a methodology for a basic eco-pedological hazard assessment* [Field paper].
- Beugler-Bell, H., Buch, M. W., & Trippner, C. (1993b). *An extended soil classification system for the Etosha National Park and adjacent areas in central northern Namibia. Based on the Revised FAO Legend of the Soil Map of the World*. (Field Paper No. 1).
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer. <https://link.springer.com/book/9780387310732>
- Bishop, T. F. A., McBratney, A. B., & Laslett, G. M. (1999). Modelling soil attribute depth functions with equal-area quadratic smoothing splines. *Geoderma*, 91(1–2), 27–45. [https://doi.org/10.1016/S0016-7061\(99\)00003-8](https://doi.org/10.1016/S0016-7061(99)00003-8)
- Bispo, A., Eglin, T., Feix, I., Le Mouel, C., Gohin, A., Forslund, A., & Levert, F. (2014). On the need of soil data for environmental assessments: The biofuel example. In D. Arrouays, N. McKenzie, J. Hempel, A. C. Richer de Forges, & A. B. McBratney (Eds), *Proceedings of the 1st GlobalSoilMap Conference, 7–9 October 2013, Orleans, France* (pp. 21–25). Taylor & Francis Group / CRC Press/Balkema.
- Bodily, J. M. (2005). *Developing a digital soil survey update protocol at the Golden Spike National Historic site* [M.S. thesis].

Utah State University.

- Boettinger, J. L., Ramsey, R. D., Bodily, J. M., Cole, N. J., Kienast-Brown, S., Nield, S. J., Saunders, A. M., & Stum, A. K. (2008). Landsat spectral data for digital soil mapping. In A. E. Hartemink, A. B. McBratney, & M. L. Mendonça Santos (Eds), *Digital Soil Mapping with Limited Data* (pp. 193–202). Springer.
- Bohn, M. P., & Miller, B. A. (2024). Locally enhanced digital soil mapping in support of a bottom-up approach is more accurate than conventional soil mapping and top-down digital soil mapping. *Geoderma*, 442, 116781. <https://doi.org/10.1016/j.geoderma.2024.116781>
- Bouma, J. (2014). Soil science contributions towards sustainable development goals and their implementation: Linking soil functions with ecosystem services. *Geoderma*, 177, 111–120. <https://doi.org/10.1002/jpln.201300646>
- Bouslih, Y., John, K., Miftah, A., Azmi, R., Aboutayeb, R., Bouasria, A., Razouk, R., & Hssaini, L. (2024). The effect of covariates on soil organic matter and pH variability: a digital soil mapping approach using random forest model. *Annals of GIS* 30(2), 215–232. <https://doi.org/10.1080/19475683.2024.2309868>
- Brady, N. C., & Weil, R. R. (2016). *The nature and properties of soils* (15th edn). Pearson. <https://doi.org/10.2136/sssaj2016.0005br>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brown, C. F., Kazmierski, M. R., Pasquarella, V. J., Rucklidge, W. J., Samsikova, M., Zhang, C., Shelhamer, E., Lahera, E., Wiles, O., Ilyushchenko, S., Gorelick, N., Zhang, L. L., Alj, S., Schechter, E., Askay, S., Guinan, O., Moore, R., Boukouvalas, A., & Kohli, P. (2025). *AlphaEarth foundations: An embedding field model for accurate and efficient global mapping from sparse label data* (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.2507.22291>
- Brungard, C. W., Boettinger, J. L., Duniway, M. C., Wills, S. A., & Edwards, T. C., Jr. (2015). Machine learning for predicting soil classes in three semi-arid landscapes. *Geoderma*, 239–240, 68–83. <https://doi.org/10.1016/j.geoderma.2014.09.019>
- Brus, D. J. (2019). Sampling for digital soil mapping: A tutorial supported by R scripts. *Geoderma*, 338, 464–480. <https://doi.org/https://doi.org/10.1016/j.geoderma.2018.07.036>
- Brus, D. J., Kempen, B., & Heuvelink, G. B. M. (2011). Sampling for validation of digital soil maps. *European Journal of Soil Science*, 62(3), 394–407. <https://doi.org/10.1111/j.1365-2389.2011.01364.x>
- Buch, M. W. (1990). *Soils, soil erosion and vegetation in the Etosha National Park / Northern Namibia—Field & laboratory results of the investigations of the year 1989. Part I — field results; Part II — laboratory results* [Project report]. University of Regensburg.
- Buch, M. W. (1993). *Känozoischer klima- und umweltwandel in Etoscha / nord-Namibia—Untersuchungen zur klimasensibilität und geomorphodynamik eines semi-ariden landschaftsraumes im südlichen Afrika—Abbildungsband* [PhD dissertation]. Regensburg University.
- Buch, M. W. (1997). Etosha pan—The third largest lake in the world? *Madoqua*, 1997(1), 49–64.
- Buch, M. W., Beugler-Bell, H., & Trippner, C. (1997). *Atlas of the project areas – northern Grootberg farms (Kunene Region) and Uuvudhiya Constituency (Oshana Region)* [Project report].
- Buch, M. W., & Trippner, C. (1997). Overview of the geological and geomorphological evolution of the Etosha region, northern Namibia. *Madoqua*, (1), 65–74.
- Buchhorn, M., Smets, B., Bertels, L., Roo, B. D., Lesiv, M., Tsendbazar, N.-E., Herold, M., & Fritz, S. (2020). *Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe* (Version V3.0.1) [Data set]. Zenodo . <https://doi.org/10.5281/ZENODO.3939050>
- Buchhorn, M., Smets, B., Bertels, L., Roo, B. D., Lesiv, M., Tsendbazar, N.-E., Li, L., & Tarko, A. (2020). *Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: product user manual* (Version Dataset v3.0, doc issue 3.3) [User manual]. Zenodo. <https://doi.org/10.5281/ZENODO.3938963>
- Buenemann, M., Coetzee, M. E., Kutuahupira, J., Maynard, J. J., & Herrick, J. E. (2023). Errors in soil maps: The need for better on-site estimates and soil map predictions. *PLoS One*, 18(1), e0270176. <https://doi.org/10.1371/journal.pone.0270176>
- Bui, E. N. (2004). Soil survey as a knowledge system. *Geoderma*, 120, 17–26. <https://doi.org/https://doi.org/10.1016/j.geoderma.2003.07.006> Get rights and content
- Bui, E. N., & Moran, C. J. (2003). A strategy to fill gaps in soil survey over large spatial extents: An example from the Murray–Darling basin of Australia. *Geoderma*, 111, 21–44.
- Bui, E. N., Searle, R. D., Wilson, P. R., Philip, S. R., Thomas, M., Brough, D., Harms, B., Hill, J. V., Holmes, K., Smolinski, H. J., & Van Gool, D. (2020). Soil surveyor knowledge in digital soil mapping and assessment in Australia. *Geoderma Regional*, 22,

- e00299. <https://doi.org/10.1016/j.geodrs.2020.e00299>
- Burrough, P. A. (1986). *Principles of geographical information systems for land resources assessment*. Taylor & Francis. <https://doi.org/https://doi.org/10.1080/10106048609354060>
- Burrough, P. A. (1989). Fuzzy mathematical methods for soil survey and land evaluation. *Journal of Soil Science*, 40(3), 477–492. <https://doi.org/10.1111/j.1365-2389.1989.tb01290.x>
- Burrough, P. A. (1991). Soil information systems. In D. J. Maguire, M. Goodchild, & D. W. Rhind (Eds), *Geographical information systems: Principles and applications* (Vol. 2, pp. 153–169). Longman Scientific and Technical.
- Burrough, P. A., Bouma, J., & Yates, S. R. (1994). The state of the art in pedometrics. *Geoderma*, 62(1–3), 311–326. [https://doi.org/10.1016/0016-7061\(94\)90043-4](https://doi.org/10.1016/0016-7061(94)90043-4)
- Calisto, L., Moreira de Sousa, L., & Batjes, N. H. (2023). *Standardised soil profile data for the world (WoSIS snapshot—December 2023)* [Data set]. ISRIC - World Soil Information (WDC - Soils). <https://doi.org/10.17027/ISRIC-WDCSOILS-20231130>
- Cambule, A. H. (2013). *Assessment of soil organic carbon stocks in the Limpopo National Park: From legacy data to digital soil mapping* [PhD thesis, ITC, University of Twente]. http://www.itc.nl/library/papers_2013/phd/cambule.pdf
- Cambule, A. H., Rossiter, D. G., Stoorvogel, J. J., & Smaling, E. M. A. (2015). Rescue and renewal of legacy soil resource inventories: A case study of the Limpopo National Park, Mozambique. *CATENA*, 125, 169–182. <https://doi.org/10.1016/j.catena.2014.10.019>
- Camps-Valls, G., Campos-Taberner, M., Moreno-Martínez, Á., Walther, S., Duveiller, G., Cescatti, A., Mahecha, M. D., Muñoz-Mari, J., García-Haro, F. J., Guanter, L., Jung, M., Gamon, J. A., Reichstein, M., & Running, S. W. (2021). A unified vegetation index for quantifying the terrestrial biosphere. *Science Advances*, 7(9), eabc7447. <https://doi.org/10.1126/sciadv.abc7447>
- Carvalho Ribeiro, E. D., Batjes, N. H., & van Oostrum, A. J. M. (2020). *World soil information service (WoSIS)—Towards the standardization and harmonization of world soil data: Procedures manual 2020*. <https://doi.org/10.17027/isric-wdc-2020-01>
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>
- Chen, S., Arrouays, D., Mulder, V. L., Poggio, L., Minasny, B., Roudier, P., Libohova, Z., Lagacherie, P., Shi, Z., Hannam, J., Meersmans, J., Richer-de-Forges, A. C., & Walter, C. (2022). Digital mapping of GlobalSoilMap soil properties at a broad scale: A review. *Geoderma*, 409, 115567. <https://doi.org/https://doi.org/10.1016/j.geoderma.2021.115567>
- Coetzee, M. E. (2001). *NamSOTER — A SOTER database for Namibia*. Ministry of Agriculture, Water and Forestry.
- Coetzee, M. E. (2009). *Chemical characterisation of the soils of east-central Namibia* [M.Sc.Agric. Thesis, University of Stellenbosch]. <http://scholar.sun.ac.za/handle/10019.1/4068>
- Coetzee, M. E. (2021). *National soil map of Namibia v1* [Map]. Agri-ecological Services, for the Ministry of Agriculture, Water and Land Reform, Republic of Namibia. Project CS/RFQ/25-2/2017: Agro-Ecological Zone and Carrying Capacity Maps.
- Coetzee, M. E. (2025). *The Namibian soil profile database (NSPD2025): An updated, expanded and harmonised compilation of national soil data* (Version 1) [Data set]. Zenodo. <https://doi.org/10.5281/ZENODO.17618737>
- Coetzee, M. E., Beernaert, F., & Calitz, A. J. (1999). Soil survey of Uitkomst Research Station. *Agricola*, 10, 17–25.
- Conrad, O. (2012). *Melton ruggedness number. Saga 9.4.1. Tool library documentation*. http://saga-gis.sourceforge.io/saga_tool_doc/9.4.1/ta_hydrology_23.html
- Cowan, D., Lebre, P., Amon, C., Becker, R., Boga, H., Boulangé, A., Chiyaka, T., Coetzee, T., De Jager, P., Dikinya, O., Eckardt, F., Greve, M., Harris, M., Hopkins, D., Houngnandan, H., Houngnandan, P., Jordaan, K., Kaimoyo, E., Kambura, A., ... Zeze, A. (2022). Biogeographical survey of soil microbiomes across sub-Saharan Africa: Structure, drivers, and predicted climate-driven changes. *Microbiome*, 10(1), 131. <https://doi.org/10.1186/s40168-022-01297-w>
- Creamer, R. E., Hagens, M., Baartman, J., & Helming, K. (2021). Editorial for special issue on “understanding soil functions – from ped to planet”. *European Journal of Soil Science*, 72(4), 1493–1496. <https://doi.org/10.1111/ejss.13099>
- Cress, J. J., Sayre, R. G., Comer, P., & Warner, H. (2009). *Terrestrial ecosystems—Land surface forms of the conterminous United States* [Data set]. USGS. <https://doi.org/10.3133/sim3085>
- Dharumarajan, S., Hegde, R., & Singh, S. K. (2017). Spatial prediction of major soil properties using random forest techniques: a case study in semi-arid tropics of south India. *Geoderma Regional*, 10, 154–162. <http://dx.doi.org/10.1016/j.geodrs.2017.07.005>
- De Blécourt, M., Röder, A., Gröngroft, A., Baumann, S., Frantz, D., & Eschenbach, A. (2018). Deforestation for agricultural expansion in SW Zambia and NE Namibia and the impacts on soil fertility, soil organic carbon- and nutrient levels. *Biodiversity*

- & *Ecology*, 6, 242–250. <https://doi.org/10.7809/b-e.00330>
- De Pauw, E. (1996). *Agro-ecological zones of Namibia. First approximation. FAO TCP/NAM/6611* [Technical Report 1]. FAO.
- De Pauw, E., & Coetzee, M. E. (1999). Production of an agro-ecological zones map of Namibia (first approximation). Part I: Condensed methodology. *Agricola*, 1998/9, 27–31.
- De Pauw, E., Coetzee, M. E., Calitz, A. J., Beukes, H., & Vits, C. (1999a). Agro-ecological zone descriptions. *Agricola*, 1998/9, 70.
- De Pauw, E., Coetzee, M. E., Calitz, A. J., Beukes, H., & Vits, C. (1999b). Production of an agro-ecological zones map of Namibia (first approximation). Part II: Results. *Agricola*, 1998/9, 33–43.
- De Sousa, L., Kempen, B., Mendes De Jesus, J., Yigini, Y., Viatkin, K., Medyckyj-Scott, D., Richie, A., Wilson, P., van Egmond, F., & Baritz, R. (2021). *Conceptual design of the global soil information system infrastructure*. FAO, ISRIC. <https://openknowledge.fao.org/handle/20.500.14283/cb4355en>
- Dent, D. L., & Ahmed, F. B. (1995). Resurrection of soil surveys: A case study of the acid sulphate soils of The Gambia. I. Data validation, taxonomic and map units. *Soil Use and Management*, 11, 69–76. <https://doi.org/10.1111/j.1475-2743.1995.tb00499.x>
- Department of Water Affairs. (1987). *Preliminary feasibility report on the irrigation possibilities along the lower Orange River*.
- Dijkshoorn, J. A. (2003). *SOTER database for southern Africa (SOTERSAF ver. 1.0)* [Technical Report]. ISRIC.
- Dobos, E., Bialkó, T., Micheli, E., & Kobza, J. (2010). Legacy soil data harmonization and database development. In J. L. Boettinger, D. W. Howell, A. C. Moore, A. E. Hartemink, & S. Kienast-Brown (Eds), *Digital Soil Mapping* (pp. 309–323). Springer Netherlands. https://doi.org/10.1007/978-90-481-8863-5_25
- Dobos, E., Daroussin, J., & Montanarella, L. (2005). *An SRTM-based procedure to delineate SOTER Terrain Units on 1:1 and 1:5 million scales*. Office for Official Publications of the European Communities.
- Dobos, E., Daroussin, J., & Montanarella, L. (2007). A quantitative procedure for building physiographic units for the European SOTER database. *Lecture Notes in Geoinformation and Cartography*, (9783540367307), 227–258. https://doi.org/10.1007/978-3-540-36731-4_10
- Dobos, E., Seres, A., Vadnai, P., Michéli, E., Fuchs, M., Láng, V., Bertóti, R. D., & Kovács, K. (2013). Soil parent material delineation using MODIS and SRTM data. *Hungarian Geographical Bulletin*, 62(2), 133–156. <https://ojs3.mtak.hu/index.php/hungeobull/article/view/2951>
- Dokuchaev, V. V. (1883). Russian chernozem. In S. Monson (Ed.), *Selected Works of V.V. Dokuchaev, 1948* (Vol. 1, pp. 14–419).
- Dominati, E., Patterson, M., & Mackay, A. (2010). A framework for classifying and quantifying the natural capital and ecosystem services of soils. *Ecological Economics*, 69(9), 1858–1868. <https://doi.org/10.1016/j.ecolecon.2010.05.002>
- Driessen, P., Deckers, J., Spaargaren, O., & Nachtergaele, F. (Eds.). (2001). *Lecture notes on the major soils of the world*. World Soil Resources Reports 94. FAO.
- Du Pisani, A. L. (2001). The idiosyncrasies of Namibian rainfall. *Agricola*, 12, 28–32.
- Du Plessis, H. M., & Dohse, T. E. (1978). *'n Ondersoek na verbrakkingsprobleme op persele K8–K13, Hardapbesproeiingskema* (Verslag No. 897/105/78). NIGB.
- Earth Resources Observation and Science (EROS) Center. (2000). *Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global* [Digital elevation data; Tiff]. U.S. Geological Survey. <https://doi.org/10.5066/F7PR7TFT>
- Eckardt, F. D., Drake, N., Goudie, A. S., White, K., & Viles, H. (2001). The role of playas in pedogenic gypsum crust formation in the Central Namib Desert: A theoretical model. *Earth Surface Processes and Landforms*, 26(11), 1177–1193. <https://doi.org/10.1002/esp.264>
- Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429246593>
- Eitel, B., & Eberle, J. (2001). Kastanozems in the Otjiwarongo Region (Namibia): Pedogenesis, associated soils, evidence for landscape degradation. *Erdkunde*, 55(1), 21–31.
- Eitel, B., Eberle, J., & Kuhn, R. (2002). Holocene environmental change in the Otjiwarongo thornbush savanna (Northern Namibia): Evidence from soils and sediments. *Catena*, 47, 43–62. [https://doi.org/10.1016/S0341-8162\(01\)00170-9](https://doi.org/10.1016/S0341-8162(01)00170-9)
- Ellis, F. (1988). *Die gronde van die Karoo. (The soils of the Karoo)* [PhD thesis]. University of Stellenbosch.
- Ellis, F. (1990). Note on soils with vesicular structure and other micro-morphological features in Karoo soils. *Proceedings 161h*

- Congress of the Soil Science Society of South Africa, Pretoria, 326–357.
- Ellis, F. (2002). *Contribution of termites to the formation of hardpans in soils of arid and semi-arid regions of South Africa*. 17th World Congress of Soil Science, Bangkok, Thailand.
- Ellis, F., & Lambrechts, J. J. N. (1994). *Dorbank, a reddish brown hardpan of South Africa: A proto-silcrete?* 15th World Congress of Soil Science, Acapulco, Mexico.
- Ellis, F., Oosthuizen, A. B., & Plessis, H. M. (1980). *'n Ondersoek ter aanvulling van verslag no. 897/105/78 oor verbrakkingsprobleme op die Hardap besproeiingskema* (Verslag No. 944/29/80). NIGB.
- Ellur, R., Ankappa, A.M., Dharumarajan, S., Puttavenkategowda, T., Nanjundegowda, T.M., Sannegowda, P.S., Pratap Mishra, A., Durin, B., & Dogančić, D. (2024). Soil quality assessment and its spatial variability in an intensively cultivated area in India. *Land* 2024(13), 970. <https://doi.org/10.3390/land13070970>
- Emadi, M., Taghizadeh-Mehrjardi, R., Cherati, A., Danesh, M., Mosavi, A., & Scholten, T. (2020). Predicting and mapping of soil organic carbon using machine learning algorithms in northern Iran. *Remote Sensing* 12(14), 2234. <https://doi.org/10.3390/rs12142234>
- ESDAC. (2008). *Global landform classification* [Digital map]. European Soil Data Centre (ESDAC), Joint Research Centre, European Commission. <https://esdac.jrc.ec.europa.eu/content/global-landform-classification>
- ESRI. (2020). *ArcGIS living atlas of the world* [GIS]. ArcGIS. <https://livingatlas.arcgis.com/en/home/>
- ESRI, & USGS. (2015). *Esri living atlas data: World ecological facets landform classes* [Digital map]. ArcGIS. <https://www.arcgis.com/home/item.html?id=cd817a746aa7437cbd72a6d39cdb4559>
- European Commission. (2020). *Soil information system for Africa (Soils4Africa). Horizon 2020 project no. 862900*. CORDIS. <https://cordis.europa.eu/project/id/862900>
- European Commission. (2024). *Soils4Africa project reporting and results*. CORDIS. <https://cordis.europa.eu/project/id/862900/reporting>
- FAO. (1990). *Guidelines for soil description* (3rd edn). Food and Agriculture Organization of the United Nations.
- FAO. (1995). *Digital soil map of the world and derived soil properties (version 3.2)* [Map]. Food and Agriculture Organization of the United Nations.
- FAO. (2003). *FAO digital soil map of the world (DSMW)* [Digital map]. Food and Agriculture Organization of the United Nations.
- FAO. (2006). *Guidelines for soil description* (4th edn). Food and Agriculture Organization of the United Nations.
- FAO. (2015). *Revised world soil charter*. Food and Agriculture Organization of the United Nations. <https://openknowledge.fao.org/handle/20.500.14283/i4965e>
- FAO. (2020a). *African soil partnership*. Food and Agriculture Organization of the United Nations. Global Soil Partnership. <https://www.fao.org/global-soil-partnership/regional-partnerships/africa/en/>
- FAO. (2020b). *INSII*. Food and Agriculture Organization of the United Nations. Global Soil Partnership. <http://www.fao.org/global-soil-partnership/pillars-action/4-information-data/insii/en/>
- FAO, & IIASA. (2023). *Harmonized world soil database version 2.0*. FAO; International Institute for Applied Systems Analysis (IIASA). <https://doi.org/10.4060/cc3823en>
- FAO/GLOSOLAN. (2019). *Standard operating procedure for soil organic carbon. Walkley–Black method. Titration and colorimetric method*. Global Soil Laboratory Network; Food and Agriculture Organization of the United Nations.
- FAO/IIASA/ISRIC/ISSCAS/JRC. (2012). *Harmonized world soil database (version 1.2)* (Version 1.2) [Database]. Food and Agriculture Organization of the United Nations; IIASA.
- FAO/ISRIC. (2003). *Soil and terrain database for southern Africa*. Food and Agriculture Organization of the United Nations.
- FAO/UNESCO. (1971). *Soil map of the world, 1:5 000 000* [Data set]. FAO; UNESCO.
- FAO/UNESCO. (1974). *Legend of the soil map of the world* [Map]. UNESCO.
- FAO/UNESCO/ISRIC. (1990). *Revised legend of the soil map of the world*. Food and Agriculture Organization of the United Nations.
- Fey M., Hughes J., Lambrechts J., Milewski A., & Mills A. (2010). *Soils of South Africa*. Cambridge University Press. <https://doi.org/https://doi.org/10.1017/CBO9780511782183>
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International*

- Journal of Climatology*, 37(12), 4302–4315. <https://doi.org/10.1002/joc.5086>
- Forkuor, G., Hounkpatin, O. K. L., Welp, G., & Thiel, M. (2017). High resolution mapping of soil properties using remote sensing variables in south-western Burkina Faso: A comparison of machine learning and multiple linear regression models. *PLOS ONE*, 12, e0170478.
- Francis, M. L. (2008). *Soil formation on the Namaqualand coastal plain* [Ph.D. thesis, Stellenbosch University]. <https://doi.org/http://scholar.sun.ac.za/handle/10019.1/1182>
- Francis, M. L., Fey, M. V., Prinsloo, H. P., Ellis, F., Mills, A. J., & Medinski, T. V. (2007). Soils of Namaqualand: Compensations for aridity. *Journal of Arid Environments*, 70(4), 588–603. <https://doi.org/https://doi.org/10.1016/j.jaridenv.2006.12.028>
- Gallant, J. C., & Wilson, J. P. (1996). TAPES-G: a grid-based terrain analysis program for the environmental sciences. *Computers & Geosciences*, 22, 713–722. [https://doi.org/https://doi.org/10.1016/0098-3004\(96\)00002-7](https://doi.org/https://doi.org/10.1016/0098-3004(96)00002-7)
- Garzanti, E., Vermeesch, P., Andò, S., Lustrino, M., Padoan, M., & Vezzoli, G. (2014). Ultra-long distance littoral transport of Orange sand and provenance of the Skeleton Coast erg (Namibia). *Marine Geology*, 357, 25–36. <https://doi.org/10.1016/j.margeo.2014.07.005>
- GDAL Development Team. (2024). GDALDEM utility documentation (slope, aspect, hillshade, TRI, TPI, roughness modes). In *GDAL/OGR documentation*. <https://gdal.org/en/stable/programs/gdaldem.html>.
- Gelsleichter, Y. (2026a). *Namibia Map* (Version 1) [JavaScript]. Google Earth Engine. https://code.earthengine.google.com/?accept_repo=users/Namibia_map/Soil_properties
- Gelsleichter, Y. (2026b). *NamSoil* [Computer software]. <https://github.com/Gelsleichter/NamSoil>
- Gelsleichter, Y., & Coetzee, M. (2026). *NamSoil v1.0 – R and Google Earth Engine (GEE) code for digital soil mapping of Namibia at 90 m resolution* (Version 1.0) [Computer software]. Zenodo. <https://doi.org/10.5281/ZENODO.18776302>
- Geological Survey of Namibia. (n.d.). *Geological map of Namibia: Digital vector compilation of the 1:250 000 map series* [Digital map]. Ministry of Mines and Energy.
- Geological Survey of Namibia. (1980). *Geological map of Namibia: 1:1,000,000* [Printed map]. Ministry of Mines and Energy.
- Gomes, L. C., Faria, R. M., de Souza, E., Veloso, G. V., Schaefer, C. E. G., & Fernandes Filho, E. I. (2019). Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma*, 340, 337–350. <https://doi.org/https://doi.org/10.1016/j.geoderma.2019.01.007>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Goudie, A. (2013). *Arid and semi-arid geomorphology*. Cambridge university press. <https://doi.org/10.1017/CBO9780511794261>
- Goudie, A., & Viles, H. (2015). *Landscapes and landforms of Namibia*. Springer Science & Business Media. <https://doi.org/https://doi.org/10.1007/978-94-017-8020-9>
- Goyal, A., & Nash, J. (2017). *Reaping richer returns: Public spending priorities for African agriculture productivity growth*. Washington, DC: World Bank and Agence Francaise de Developpement. <https://doi.org/10.1596/978-1-4648-0937-8>
- Gray, J. M., Bishop, T. F. A., & Wilford, J. R. (2016). Lithology and soil relationships for soil modelling and mapping. *Catena*, 147(October 2017), 429–440. <https://doi.org/10.1016/j.catena.2016.07.045>
- Greiner, L., Keller, A., Grêt-Regamey, A., & Papritz, A. (2017). Soil function assessment: Review of methods for quantifying the contributions of soils to ecosystem services. *Land Use Policy*, 69, 224–237. <https://doi.org/10.1016/j.landusepol.2017.06.025>
- Grimm, R., Behrens, T., Märker, M., & Elsenbeer, H. (2008). Soil organic carbon concentrations and stocks on Barro Colorado Island—Digital soil mapping using Random Forests analysis. *Geoderma*, 146(1), 102–113. <https://doi.org/https://doi.org/10.1016/j.geoderma.2008.05.008>
- Grunwald, S. (2009). Multi-criteria characterization of recent digital soil mapping and modeling approaches. *Geoderma*, 152(3–4), 195–207. <https://doi.org/10.1016/j.geoderma.2009.06.003>
- Grunwald, S., Thompson, J. A., & Boettinger, J. L. (2011). Digital soil mapping and modeling at continental scales: Finding solutions for global issues. *Soil Science Society of America Journal*, 75(4), 1201–1213. <https://doi.org/10.2136/sssaj2011.0025>
- Guo, P. T., Li, M. F., Luo, W., Tang, Q. F., Liu, Z. W., & Lin, Z. M. (2015). Digital mapping of soil organic matter for rubber plantation at regional scale: An application of random forest plus residuals kriging approach. *Geoderma*, 237–238, 49–59. <https://doi.org/10.1016/J.GEODERMA.2014.08.009>
- Guo, J., Wang, K., & Jin, S. (2022). Mapping of soil pH based on SVM-RFE feature selection algorithm. *Agronomy* 12(11), 2742. <https://doi.org/10.3390/agronomy12112742>

- Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, 46(1–3), 389–422. <https://doi.org/10.1023/A:1012487302797>
- Haddon, I. G., & McCarthy, T. S. (2005). The Mesozoic–Cenozoic interior sag basins of central Africa: The Late-Cretaceous–Cenozoic Kalahari and Okavango basins. *Journal of African Earth Sciences*, 43(1–3), 316–333. <https://doi.org/10.1016/j.jafrearsci.2005.07.008>
- Han, S. Y., Filippi, P., Singh, K., Whelan, B. M., & Bishop, T. F. A. (2022). Assessment of global, national and regional-level digital soil mapping products at different spatial supports. *European Journal of Soil Science*, 73(5), e13300. <https://doi.org/10.1111/ejss.13300>
- Hartemink, A. E. (2008). Soils are back on the global agenda. *Soil Use and Management*, 24(4), 327–330. <https://doi.org/10.1111/j.1475-2743.2008.00187.x>
- Hartemink, A. E., & McBratney, Alex. (2008). A soil science renaissance. *Geoderma*, 148(2), 123–129. <https://doi.org/10.1016/j.geoderma.2008.10.006>
- Hartemink, A. E., & Minasny, B. (2014). Towards digital soil morphometrics. *Geoderma*, 230–231, 305–317.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning*. Springer.
- Heine, K., & Völkel, J. (2010). Soil clay minerals in Namibia and their significance for the terrestrial and marine past global change research. *African Study Monographs, Supplement*, pp. 31–50.
- Heine, K., & Walter, R. (1996). Gypcretes of the central Namib desert, Namibia, *Palaeoecology of Africa*, 24, pp. 173–201.
- Hengl, T., Rossiter, D. G., & Stein, A. (2003). Soil sampling strategies for spatial prediction by correlation with auxiliary maps. *Australian Journal of Soil Research*, 41, 1403–1422. <https://doi.org/10.1071/SR03005>
- Hengl, T., De Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B. M., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J. G. B., Walsh, M. G., & Gonzalez, M. R. (2014). SoilGrids1km—Global soil information based on automated mapping. *PLoS ONE*, 9(8), e105992. <https://doi.org/10.1371/journal.pone.0105992>
- Hengl, T., Heuvelink, G. B. M., Kempen, B., Leenaars, J. G. B., Walsh, M. G., Shepherd, K. D., Sila, A., MacMillan, R. A., Mendes De Jesus, J., Tamene, L., & Tondoh, J. E. (2015). Mapping soil properties of Africa at 250 m resolution: Random Forests significantly improve current predictions. *PLOS ONE*, 10(6), e0125814. <https://doi.org/10.1371/journal.pone.0125814>
- Hengl, T., Leenaars, J. G. B., Shepherd, K. D., Walsh, M. G., Heuvelink, G. B. M., Mamo, T., Tilahun, H., Berkhout, E., Cooper, M., Fegraus, E., Wheeler, I., & Kwabena, N. A. (2017a). Soil nutrient maps of sub-Saharan Africa: Assessment of soil nutrient content at 250 m spatial resolution using machine learning. *Nutrient Cycling in Agroecosystems*, 109(1), 77–102. <https://doi.org/10.1007/s10705-017-9870-x>
- Hengl, T., Mendes De Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017b). SoilGrids250m: Global gridded soil information based on machine learning. *PLOS ONE*, 12(2), e0169748. <https://doi.org/10.1371/journal.pone.0169748>
- Hengl, T., Miller, M. A. E., Krizán, J., Shepherd, K. D., Sila, A., Kilibarda, M., Antonijević, O., Glušica, L., Dobermann, A., Haefele, S. M., McGrath, S. P., Acquah, G. E., Collinson, J., Parente, L., Sheykhmousa, M., Saito, K., Johnson, J.-M., Chamberlin, J., Silatsa, F. B. T., ... Crouch, J. (2021). African soil properties and nutrients mapped at 30 m spatial resolution using two-scale ensemble machine learning. *Scientific Reports*, 11(1), 6130. <https://doi.org/10.1038/s41598-021-85639-y>
- Herrick, J. E., Beh, A., Barrios, E., Bouvier, I., Coetzee, M., Dent, D., Elias, E., Hengl, T., Karl, J. W., Liniger, H., Matuszak, J., Neff, J. C., Ndungu, L. W., Obersteiner, M., Shepherd, K. D., Urama, K. C., Bosch, R., & Webb, N. P. (2016). The land-potential knowledge system (LandPKS): Mobile apps and collaboration for optimizing climate change investments. *Ecosystem Health and Sustainability*, 2(3), e01209. <https://doi.org/10.1002/ehs2.1209>
- Heung, B., Bulmer, C. E., & Schmidt, M. G. (2014). Predictive soil parent material mapping at a regional-scale: A random forest approach. *Geoderma*, 214–215, 141–154.
- Heung, B., Ho, H. C., Zhang, J., Knudby, A., Bulmer, C. E., & Schmidt, M. G. (2016). An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. *Geoderma*, 265, 62–77. <https://doi.org/10.1016/j.geoderma.2015.11.014>
- Heuvelink, G. B. M. (2003). The definition of pedometrics. *Pedometron*, 15, 11–12. <https://doi.org/http://pedometrics.org/pedometron/>
- Heuvelink, G. B. M. (2018). Uncertainty and uncertainty propagation in soil mapping and modelling. In Alex. B. McBratney, B. Minasny, & U. Stockmann (Eds), *Pedometrics* (pp. 439–461). Springer International Publishing. <https://doi.org/10.1007/978->

- Hiederer, R., Joint Research Centre, & Institute for Environment and Sustainability. (2013). *Mapping soil typologies: Spatial decision support applied to the European Soil Database*. Office for Official Publications of the European Communities. <https://doi.org/10.2788/87286>
- Hijmans, R. J. (2025). *Terra: Spatial data analysis. R package* (Version 1.8–60) [Computer software]. <https://CRAN.R-project.org/package=terra>
- Hobson, R. D. (1972). Surface roughness in topography: Quantitative approach. In R. J. Chorley (Ed.), *Spatial Analysis in Geomorphology* (pp. 221–245). Harper & Row/ Methuen & Co.
- Hudson, B. D. (1992). The soil survey as paradigm-based science. *Soil Science Society of America Journal*, 56, 836–841.
- ICC & MAWRD. (2000). *Project to support the Agro-Ecological Zoning Programme (AEZ) in Namibia*. Institut Cartogràfic de Catalunya, Ministry of Agriculture Water and Rural Development.
- IIASA, & FAO. (2012). *Global agro-ecological zones (GAEZ v3.0)*. IIASA; FAO.
- ISRIC–WSI. (n.d.). *World soil information service (WoSIS) – latest* [Dashboard]. ISRIC Soil Data Hub. WoSIS. Retrieved https://dashboards.isric.org/superset/dashboard/wosis_latest/
- ISRIC–WSI. (2023). *Africa soil information system (AfricaSIS)* [Data set]. <https://africasis.isric.org/>
- ISSS/ISRIC/FAO. (1998). *World reference base for soil resources* (1st edn). International Society of Soil Science (ISSS); International Soil Reference and Information Centre (ISRIC); Food and Agriculture Organisation of the United Nations (FAO).
- IUSS Working Group WRB (Ed.). (2015). *World reference base for soil resources 2014: International soil classification system for naming soils and creating legends for soil maps. Update 2015*. FAO. <http://www.fao.org/3/i3794en/i3794en.pdf>
- Iwahashi, J., Kamiya, I., Matsuoka, M., & Yamazaki, D. (2018). Global terrain classification using 280 m DEMs: Segmentation, clustering, and reclassification. *Progress in Earth and Planetary Science*, 5(1), 1. <https://doi.org/10.1186/s40645-017-0157-2>
- Iwahashi, J., & Pike, R. J. (2007). Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. *Geomorphology*, 86(3–4), 409–440.
- IWRMP Joint Venture Namibia. (2010). *Integrated water resources management plan for Namibia* [Management plan]. Ministry of Agriculture, Water and Forestry.
- Jenness, J. (2002). *Surface areas and ratios from elevation grid (surfgrids.avs) extension for ArcView 3.x* (Version 1.2) [Computer software]. Jenness Enterprises. http://www.jennessent.com/arcview/surface_areas.htm
- Jenness, J. (2006). *Topographic position index (tpi_jen.avx) extension for ArcView 3.x* (Version 1.3a) [Computer software]. <http://www.jennessent.com/arcview/tpi.htm>
- Jenness, J. (2013). *Manual: DEM surface tools for ArcGIS*. Jenness Enterprises.
- Jenny, H. (1941). *Factors of soil formation: A system of quantitative pedology*. McGraw-Hill.
- Jenny, H., Salem, A. E., & Wallis, J. R. (1968). Interplay of soil organic matter and soil fertility with state factors and soil properties. *Pontifical Academy of Sciences, Scripta Varia*, 32, 5–36.
- Jensen, J. R. (2005). *Introductory digital image processing: A remote sensing perspective* (3rd edn). Pearson Prentice Hall.
- Jones, A., Breuning-Madsen, H., Brossard, M., Dampha, A., Deckers, J., Dewitte, O., Gallali, T., Hallett, S., Jones, R., Kilasara, M., Le Roux, P., Micheli, E., Montanarella, L., Spaargaren, O., Thiombiano, L., Ranst, E., Yemefack, M., & Zougmore R. (2013). *Soil atlas of Africa*. European Commission, Publications Office of the European Union.
- Kasraei, B., Schmidt, M. G., Zhang, J., Bulmer, C. E., Filatow, D. S., Arbor, A., Pennell, T., & Heung, B. (2024). A framework for optimizing environmental covariates to support model interpretability in digital soil mapping. *Geoderma*, 445, 116873. <https://doi.org/10.1016/j.geoderma.2024.116873>
- Kassai, P., Kocsis, M., Szatmári, G., Makó, A., Mészáros, J., Laborczy, A., Magyar, Z., Takács, K., Pásztor, L., & Szabó, B. (2025). Large-scale mapping of soil particle size distribution using legacy data and machine learning-based pedotransfer functions. *Geoderma*, 454, 117178. <https://doi.org/10.1016/j.geoderma.2025.117178>
- Keith, D. A., Ferrer-Paris, J. R., Nicholson, E., Bishop, M. J., Polidoro, B. A., Ramirez-Llodra, E., Tozer, M. G., Nel, J. L., Mac Nally, R., Gregr, E. J., Watermeyer, K. E., Essl, F., Faber-Langendoen, D., Franklin, J., Lehmann, C. E. R., Etter, A., Roux, D. J., Stark, J. S., Rowland, J. A., ... Kingsford, R. T. (2022). A function-based typology for Earth's ecosystems. *Nature*, 610(7932), 513–518. <https://doi.org/10.1038/s41586-022-05318-4>
- Kempen, B., Brus, D. J., Stoorvogel, J. J., Heuvelink, G. B. M., & De Vries, F. (2012). Efficiency comparison of conventional and

- digital soil mapping for updating soil maps. *Soil Science Society of America Journal*, 76(6), 2097–2115. <https://doi.org/10.2136/sssaj2011.0424>
- Kempf, J. (1999). *Zur geomorphogenese und pedogenese zentral-Namibias* [PhD thesis]. Würzburg.
- Kienast-Brown, S., Libohova, Z., & Boettinger, J. (2017). Digital soil mapping. In Soil Science Division Staff (Ed.), *Soil Survey Manual*. United States Department of Agriculture.
- Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5). <https://doi.org/10.18637/jss.v028.i05>
- Kuhn, M. (2024). *Caret* (Version 7.0–1) [Computer software]. <https://rdocumentation.org/packages/caret/versions/7.0-1>
- Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer New York. <https://doi.org/10.1007/978-1-4614-6849-3>
- Kumar, A., Moharana, P. C., Jena, R. K., Malyan, S. K., Sharma, G. K., Fagodiya, R. K., Shabnam, A. A., Jigyasu, D. K., Kumari, K. M. V., & Doss, S. G. (2023). Digital mapping of soil organic carbon using machine learning algorithms in the upper Brahmaputra valley of northeastern India. *Land*, 12(10), 1841. <https://doi.org/10.3390/land12101841>
- Kumaraperumal, R., Pazhanivelan, S., Geethalakshmi, V., Nivas Raj, M., Muthumanickam, D., Kaliaperumal, R., Shankar, V., Nair, A.M., Yadav, M.K., & Tarun Kshatriya, T.V. (2022). Comparison of machine learning-based prediction of qualitative and quantitative digital soil-mapping approaches for eastern districts of Tamil Nadu, India. *Land* 2022(11), 2279. <https://doi.org/10.3390/land11122279>
- Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the boruta package. *Journal of Statistical Software*, 36(11). <https://doi.org/10.18637/jss.v036.i11>
- Kutuahupira, J. K., & Mouton, H. D. (2006a). *Soils report: Auas-Oanob Conservancy* [Survey report]. Agro-ecological Zoning Programme, Ministry of Agriculture, Water and Forestry.
- Kutuahupira, J. K., & Mouton, H. D. (2006b). *Soils report: Quantification of land production potential project* [Survey report]. Agro-ecological Zoning Programme, Ministry of Agriculture, Water and Forestry.
- Kutuahupira, J. T., Mouton, H. D., & Beukes, H. A. (2003). *Soil survey of Otjinene, Epukiro, Gam and Okakarara* [Survey report]. Agro-ecological Zoning Programme, Ministry of Agriculture, Water and Forestry.
- Kutuahupira, J. T., Mouton, H. G., & Coetzee, M. E. (2001a). Soil survey of Mahenene Research Station. *Agricola*, 12, 59–66.
- Kutuahupira, J. T., Mouton, H. G., & Coetzee, M. E. (2001b). Soil survey of Sonop Research Station. *Agricola*, 12, 67–72.
- Lagacherie, P., & McBratney, A. B. (2007). Spatial soil information systems and spatial soil inference systems: Perspectives for digital soil mapping. In P. Lagacherie, A. B. McBratney, & M. Voltz (Eds), *Digital Soil Mapping: An introductory perspective* (pp. 3–24). Elsevier.
- Lambrechts, J. J. N. (1977). *Verslag oor 'n grondopname vir besproeiing in die weste van Ovambo* [Survey report].
- Lambrechts, J. J. N., & MacVicar, C. N. (2004). Soil genesis and classification and soil resources databases. *South African Journal of Plant and Soil*, 21(5), 288–300. <https://doi.org/10.1080/02571862.2004.10635065>
- Lambrechts, J. J. N., & Schloms, B. (1993). *Reconnaissance soil survey of a section of Tandjieskoppe, north-west of Noordoewer along the Orange River* [Survey report].
- Lamichhane, S., Kumar, L., & Wilson, B. (2019). Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review. *Geoderma*, 352, 395–413. <https://doi.org/10.1016/j.geoderma.2019.05.031>
- Láng, V., Fuchs, M., Szegi, T., Csorba, Á., & Michéli, E. (2016). Deriving world reference base reference soil groups from the prospective global soil map product—A case study on major soil types of Africa. *Geoderma*, 263, 226–233. <https://doi.org/10.1016/j.geoderma.2015.07.005>
- Leenaars, J. G. B. (2012). *Africa soil profiles database, version 1.0: A compilation of georeferenced and standardised legacy soil profile data for sub-Saharan Africa (with dataset)* (No. 2012/03; Issue 2012/03)). ISRIC. <https://edepot.wur.nl/207059>
- Leenaars, J. G. B., van Oostrum, A. J. M., & Ruiperez Gonzalez, M. (2014). *Africa soil profiles database, version 1.2. A compilation of georeferenced and standardised legacy soil profile data for Sub-Saharan Africa (with dataset)* (No. 2014/01). Africa Soil Information Service (AfSIS) project; ISRIC–World Soil Information. <https://library.wur.nl/WebQuery/isric/2259472>.
- Legates, D. R., & McCabe, G. J. (1999). Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, 35(1), 233–241.
- Lehner, B., & Grill, G. (2013). Global river hydrography and network routing: Baseline data and new approaches to study the world’s large river systems. *Hydrological Processes*, 27(15), 2171–2186. <https://doi.org/10.1002/hyp.9740>

- Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, 89(10), 93–94. <https://doi.org/10.1029/2008EO100001>
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3), 18–22.
- Lin, L. I.-K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 45(1), 255 .
<https://doi.org/10.2307/2532051>
- Liu, Y., Wu, X., Wu, T., Zhao, L., Li, R., Li, W., Hu, G., Zou, D., Ni, J., Du, Y., Wang, M., Li, Z., Wei, X., & Yan, X. (2022). Soil texture and its relationship with environmental factors on the Qinghai–Tibet plateau. *Remote Sensing*, 14(15), 3797.
<https://doi.org/10.3390/rs14153797>
- Lopes, M. E. (2019). Estimating the algorithmic variance of randomized ensembles via the bootstrap. *The Annals of Statistics*, 47(2). <https://doi.org/10.1214/18-AOS1707>
- Louw, P. A. (1957). *Die bodemopname van die Hardap-besproeiingsprojek, Mariental, SWA* (No. 500/333/57). NIGB.
- Loxton, R. F., Hunting, & Associates. (1970). *Report on a reconnaissance survey of the soils of northern South West Africa in terms of their potential for irrigation*.
- Loxton, R. F., Hunting, & Associates. (1974a). *Report on a reconnaissance survey of the natural resources of Damaraland*.
- Loxton, R. F., Hunting, & Associates. (1974b). *Report on a reconnaissance survey of the natural resources of Kaokoland*.
- Ma, Y., Minasny, B., Malone, B. P., & Mcbratney, A. B. (2019). Pedology and digital soil mapping (DSM). *European Journal of Soil Science*, 70(2), 216–235. <https://doi.org/10.1111/ejss.12790>
- MacMillan, R. A. (2008). Experiences with applied DSM: Protocol, availability, quality and capacity building. In A. E. Hartemink, A. B. McBratney, & M. L. M. Santos (Eds), *Digital soil mapping with limited data* (pp. 113–136). Springer.
https://doi.org/https://doi.org/10.1007/978-1-4020-8592-5_10
- MacMillan, R. A., Pettapiece, W. W., & Brierley, J. A. (2005). An expert system for allocating soils to landforms through the application of soil survey tacit knowledge. *Canadian Journal of Soil Science*, 85(1), 103–112. <https://doi.org/10.4141/S04-029>
- MacVicar, C. N., Villiers, J. M., Loxton, R. F., Verster, E., Lambrechts, J. J. N., Merryweather, F. R., Roux, J., Van Rooyen, T. H., & Harmse, H. J. von M. (1977). *Soil classification: A binomial system for South Africa*. Department of Agricultural Technical Services.
- Mallavan, B. P., Minasny, B., & McBratney, A. B. (2010). Homosoil, a methodology for quantitative extrapolation of soil information across the globe. In J. L. Boettinger, D. W. Howell, A. C. Moore, A. E. Hartemink, & S. Kienast-Brown (Eds), *Digital Soil Mapping Bridging Research Environmental Application, and Operation* (pp. 137–150). Springer. Heidelberg.
- Malone, B. P. (2023). *ithir: Soil data and some useful associated functions* (Version 1.0) [Computer software].
<https://github.com/brendo1001/ithir>
- Malone, B. P., McBratney, A. B., & Minasny, B. (2011). Empirical estimates of uncertainty for mapping continuous depth functions of soil attributes. *Geoderma*, 160(3–4), 614–626. <https://doi.org/10.1016/j.geoderma.2010.11.013>
- Malone, B. P., McBratney, A. B., Minasny, B., & Laslett, G. M. (2009). Mapping continuous depth functions of soil carbon storage and available water capacity. *Geoderma*, 154(1–2), 138–152. <https://doi.org/10.1016/j.geoderma.2009.10.007>
- Malone, B. P., Minasny, B., & McBratney, A. B. (2017). *Using R for digital soil mapping*. Springer International Publishing.
<https://doi.org/10.1007/978-3-319-44327-0>
- MAWRD. (1995). *Proposed new government irrigation farm on a portion of the farm Aussenkehr 147. Vol. 2 Soils Report*. Ministry of Agriculture Water and Rural Development.
- MAWRD-AEZ. (2009). *Namibian agricultural resources information system (NARIS)* [Data set]. Agro-ecological Zoning Programme, Ministry of Agriculture, Water and Forestry.
- McBratney, A. B., Evangelista, S. J., Pachón Maldonado, J. C., Styc, Q., Francos, N., Sharififar, A., & Padarian, J. (2025). The evolution of (digital) pedology – drawing from the past, adapting to the future. *Geoderma*, 463, 117588.
<https://doi.org/10.1016/j.geoderma.2025.117588>
- McBratney, A. B., Mendonça Santos, M. L., & Minasny, B. (2003). On digital soil mapping. *Geoderma*, 117(1–2), 3–52.
[https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4)
- McKenzie, N. J., & Ryan, P. J. (1999). Spatial prediction of soil properties using environmental correlation. *Geoderma*, 89, 67–94.
- MEFT. (2024). *Review and consolidation of Namibia's bush thickening maps. Final report on Bush Maps project*. Perivoli Rangeland Institute; Ministry of Environment, Forestry and Tourism; Bush Control and Biomass Utilisation Project, Deutsche

- Mendelsohn, J., & El Obeid, S. (2003). *Sand and water: A profile of the Kavango Region*. Struik Publishers; RAISON.
- Mendelsohn, J., & El Obeid, S. (2004). *Okavango River: The flow of a lifeline* (1st ed). Struik Publishers.
- Mendelsohn, J., El Obeid, S., & Roberts, C. (2000). *An environmental profile of north-central Namibia*. Ministry of Environment and Tourism.
- Mendelsohn, J., Jarvis, A., Roberts, C., & Robertson, T. (2002). *Atlas of Namibia: A portrait of the land and its people*. David Philip Publishers.
- Mendelsohn, J., & Roberts, C. (1997). *An environmental profile and atlas of Caprivi*. Ministry of Environment and Tourism & Gamsberg Macmillan.
- Meybeck, M., Green, P., & Vorosmarty, C. J. (2001). A new typology for mountains and other relief classes: An application to global continental water resources and population distribution. *Mountain Research and Development*, 21, 34–45.
- Mgohele, R. N., Massawe, B. H. J., Shitindi, M. J., Sanga, H. G., & Omar, M. M. (2024). Prediction of soil texture using remote sensing data. A systematic review. *Frontiers in Remote Sensing*, 5, 1461537. <https://doi.org/10.3389/frsen.2024.1461537>
- Miller, M. A. E., Shepherd, K. D., Kisu, B., & Collinson, J. (2021). iSDAsoil: The first continent-scale soil property map at 30 m resolution provides a soil information revolution for Africa. *PLOS Biology*, 19(11), e3001441. <https://doi.org/10.1371/journal.pbio.3001441>
- Miller, R. M. (2008). Upper Palaeozoic to Cenozoic. Namib Group: Late Tertiary to Quaternary deposits. In *The geology of Namibia* (Vol. 3). Geological Survey of Namibia, Ministry of Mines and Energy.
- Minai, J. O., Schulze, D. G., & Libohova, Z. (2022). Renewal of archival legacy soil data: A case study of the Busia area, Kenya. *Frontiers in Soil Science*, 1, 765248. <https://doi.org/10.3389/fsoil.2021.765248>
- Minasny, B., & McBratney, Alex. B. (2016). Digital soil mapping: A brief history and some lessons. *Geoderma*, 264, 301–311. <https://doi.org/10.1016/j.geoderma.2015.07.017>
- Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill.
- Monteith, J. L. (1965). Evapotranspiration and the environment. *The State and Movement of Water in Living Organisms*, 205–234.
- Monteith, J. L. (1981). Evapotranspiration and surface temperature. *Quarterly Journal Royal Meteorological Society*, 107, 1–27.
- Moriasi, D. N., Arnold, J. G., Liew, M. W. V., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3), 885–900. <https://doi.org/10.13031/2013.23153>
- Mulder, V. L., De Bruin, S., Shaepman, M. E., & Mayr, T. R. (2011). The use of remote sensing in soil and terrain mapping—A review. *Geoderma*, 162, 1–19.
- Mulder, V. L., Lacoste, M., Richer-de-Forges, A. C., Martin, M. P., & Arrouays, D. (2016). National versus global modelling the 3D distribution of soil organic carbon in mainland France. *Geoderma*, 263, 16–34.
- Murray, Biesenbach & Badenhorst Inc., Windhoek Consulting Engineers, & Namibia Development Corporation. (1997). *Identification and prioritization of irrigation development opportunities on the Orange River*. (Technical Report Nos 2010–03).
- Nachtergaele, F., van Velthuyzen H., Verelst, L., & Wiberg, D. (2012). *Manual harmonized world soil database v1.2* [Manual].
- NASA JPL. (2020). *NASADEM merged DEM global 1 arc second* (Version V001) [Data set]. NASA Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MEASURES/NASADEM/NASADEM_HGT.001
- Navarro-Hasse, E., Yáñez, C., Neaman, A., & Pinochet, D. (2023). The effects of parent rock on soil clay mineralogy and soil physicochemical properties: A review. *Idesia (Arica)*, 41(4), 125–139. <https://doi.org/10.4067/S0718-34292023000400125>
- Nellemann, C., & Fry, G. (1995). Quantitative analysis of terrain ruggedness in reindeer winter grounds. *Arctic*, 48(2), 172–176. <https://doi.org/http://www.jstor.org/stable/40511641>
- Nenkam, A. M., Wadoux, A. M. J. C., Minasny, B., McBratney, A. B., Traoré, P. C. S., & Whitbread, A. M. (2023). Using homosols to enrich sparse soil data infrastructure: An example from Mali. *Catena*, 223, 106862.
- Nenkam, A. M., Wadoux, A. M. J.-C., Minasny, B., Silatsa, F. B. T., Yemefack, M., Ugbaje, S. U., Akpa, S., Zijl, G. V., Bouasria, A., Bouslihim, Y., Chabala, L. M., Ali, A., & McBratney, A. B. (2024). Applications and challenges of digital soil mapping in Africa. *Geoderma*, 449, 117007. <https://doi.org/10.1016/j.geoderma.2024.117007>

- Nield, S. J., Boettinger, J. L., & Ramsey, R. D. (2007). Digitally mapping gypsic and natric soil areas using Landsat ETM data. *Soil Science Society of America Journal*, 71(1), 245–252.
- Nijbroek, R., Piikki, K., Söderström, M., Kempen, B., Turner, K., Hengari, S., & Mutua, J. (2018). Soil organic carbon baselines for land degradation neutrality: Map accuracy and cost tradeoffs with respect to complexity in Otjozondjupa, Namibia. *Sustainability*, 10(5), 1610. <https://doi.org/10.3390/su10051610>
- Nussbaum, M., Spiess, K., Baltensweiler, A., Grob, U., Keller, A., Greiner, L., Schaepman, M. E., & Papritz, A. (2018). Evaluation of digital soil mapping approaches with large sets of environmental covariates. *SOIL*, 4(1), 1–22. <https://doi.org/10.5194/soil-4-1-2018>
- Odeh, I. O. A., Leenaars, J. G. B., Hartemink, A. E., & Amapu, I. (2012). The challenges of collating legacy data for digital mapping of Nigerian soils. In B. Minasny, B. Malone, & A. B. McBratney (Eds), *Digital Soil Assessments and Beyond*. Taylor & Francis Group.
- Omuto, C. T., Nachtergaele, F. O., & Rojas, R. V. (2012). *State of the art report on global soil information: Where are we? Where to go?* Food and Agriculture Organization of the United Nations.
- Omuto, C. T., Weiss, L., & Wole, F. (2024). Developing soil information systems for Africa. *Soil Information Systems for Africa*. <https://bookdown.org/cthine/bookdown-sis/>
- Opperman, D. P. J., Viljoen, M. F., Toit, D., & Wilke, P. I. (Eds). (1982). *Landbou-ontwikkelingsplan vir die Oos-Capri* [Development plan].
- Padarian, J., Minasny, B., & McBratney, A. B. (2015). Using Google’s cloud-based platform for digital soil mapping. *Computers & Geosciences*, 83, 80–88.
- Padarian, J., Minasny, B., & McBratney, A.B. (2017). Chile and the Chilean soil grid: A contribution to GlobalSoilMap. *Geoderma Regional* 9, 17–28. <https://doi.org/10.1016/j.geodrs.2016.12.001>
- Page, D. (1976). *Evaluasie van hulpborne van Kaokoland en ontwikkelingsvoorstelle. Beplanningsverslag no. 1* [Development plan].
- Page, D. (1978). *Nasionale plan Kavango ('n voorlopige benadering)* [Development plan].
- Page, D. (1979). *n Raamwerk vir ontwikkeling van Kavango: Atlas* (Development plan No. 22).
- Page, D. (1980). *N Raamwerk vir ontwikkeling van Kavango* (Development plan No. 28).
- Partridge, T. C. (1987a). *Report on a soil survey in areas which could be irrigated from the proposed Neckartal Dam on the Fish River. Vol. 2 Soil profile records and laboratory results* [Survey report]. Department of Water Affairs, South West Africa.
- Partridge, T. C. (1987b). *Report on a soil survey in areas which could be irrigated from the proposed Neckartal Dam on the Fish River. Vol. Agroclimatic summary* [Survey report]. Department of Water Affairs, South West Africa.
- Pásztor, L., Dobos, E., Szatmári, G., Laborcz, A., Takács, K., Bakacsi, Zs., & Szabó, J. (2014). Application of legacy soil data in digital soil mapping for the elaboration of novel, countrywide maps of soil conditions. *Agrokémia És Talajtan*, 63(1), 79–88. <https://doi.org/10.1556/agrokem.63.2014.1.9>
- Paterson, G. (2021). *Guidance for the laboratory analysis—Soils4Africa Project* [Procedures manual]. <https://www.soils4africa-h2020.eu/serverspecific/soils4africa/images/Documents/GuidanceonLaboratoryAnalysis.pdf>
- Paterson, G., Turner, D., Wiese, L., Van Zijl, G., Clarke, C., & Van Tol, J. (2015). Spatial soil information in South Africa: Situational analysis, limitations and challenges. *South African Journal of Science*, 111(5/6), 1–7. <https://doi.org/10.17159/sajs.2015/20140178>
- Pauling, R. C., Speidel, S. E., Thomas, M. G., Holt, T. N., & Enns, R. M. (2018). Evaluation of moderate to high elevation effects on pulmonary arterial pressure measures in Angus cattle. *Journal of Animal Science*, 96(9), 3599–3605.
- Petersen, A. (2008). *Pedodiversity of southern African drylands* [PhD dissertation.]. University of Hamburg.
- Poggio, L., De Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *SOIL*, 7(1), 217–240. <https://doi.org/10.5194/soil-7-217-2021>
- Ponce-Hernandez, R., Marriott, F. H. C., & Beckett, P. H. T. (1986). An improved method for reconstructing a soil-profile from analysis of a small number of samples. *Journal of Soil Science*, 37, 455–467.
- Probst, P., Boulesteix, A.-L., & Bischl, B. (2019). Tunability: Importance of hyperparameters of machine learning algorithms. *Journal of Machine Learning Research*, 20, 1–32.

- Probst, P., Wright, M. N., & Boulesteix, A. (2019). Hyperparameters and tuning strategies for random forest. *WIREs Data Mining and Knowledge Discovery*, 9(3), e1301. <https://doi.org/10.1002/widm.1301>
- Pröpper, M., Gröngroft, A., Finckh, M., Stirn, S., De Cauwer, V., Lages, F., Masamba, W., Murray-Hudson, M., Schmidt, L., Strohbach, B., & Jürgens, N. (Eds). (2015). *The Future Okavango—Findings, scenarios and recommendations for action. Research project final synthesis. 2010–2015.* (p. 190).
- QGIS Development Team. (2024a). *QGIS geographic information system* (Version 3.40) [Computer software]. <https://www.qgis.org>.
- QGIS Development Team. (2024b). *QGIS user manual: Raster analysis (slope, aspect, hillshade, TRI, TPI, roughness)* [User manual].
- R Core Team. (2024a). *R: A language and environment for statistical computing* [Manual]. R Foundation for Statistical Computing.
- R Core Team. (2024b). *The R project for statistical computing* [Computer software]. R Foundation for Statistical Computing. www.R-project.org
- Rabot, E., Wiesmeier, M., Schlüter, S., & Vogel, H.-J. (2018). Soil structure as an indicator of soil functions: A review. *Geoderma*, 314, 122–137. <https://doi.org/10.1016/j.geoderma.2017.11.009>
- Rajendran, S., & Nasir, S. (2019). ASTER capability in mapping of mineral resources of arid region: A review on mapping of mineral resources of the Sultanate of Oman. *Ore Geology Reviews*, 108, 33–53. <https://doi.org/10.1016/j.oregeorev.2018.04.014>
- Ramcharan, A., Hengl, T., Nauman, T., Brungard, C., Waltman, S., Wills, S., & Thompson, J. (2017). Soil property and class maps of the conterminous United States at 100-meter spatial resolution. *Soil Science Society of America Journal* 82(1), 186–201. <https://doi.org/10.2136/sssaj2017.04.0122>
- Rasaei, Z., Rossiter, D. G., & Farshad, A. (2020). Rescue and renewal of legacy soil resource inventories in Iran as an input to digital soil mapping. *Geoderma Regional*, 21, e00262. <https://doi.org/10.1016/j.geodrs.2020.e00262>
- Reddy, N.N., Chakraborty, P., Roy, S., Singh, K., Minasny, B., McBratney, A.B., Biswas, A., & Das, B.S. (2021). Legacy data-based national-scale digital mapping of key soil properties in India. *Geoderma* 381, 114684. <https://doi.org/10.1016/j.geoderma.2020.114684>
- Republic of Namibia. (2007, December 27). *Environmental Management Act, No. 7 of 2007. Government Notice No. 232, Government Gazette of the Republic of Namibia No. 3966.*
- Republic of Namibia. (2011). *National policy on climate change for Namibia*. Ministry of Environment & Tourism.
- Republic of Namibia. (2012). *National rangeland management policy & strategy*. Ministry of Agriculture, Water and Forestry.
- Republic of Namibia. (2014a). *Namibia's second national biodiversity strategy and action plan—2013–2022*. Ministry of Environment & Tourism.
- Republic of Namibia. (2014b). *National climate change strategy & action plan: 2013–2020*. Ministry of Environment & Tourism.
- Republic of Namibia. (2014c). *Third national action programme for Namibia to implement the United Nations Convention to Combat Desertification: 2014–2024*. Ministry of Environment & Tourism.
- Riley, S. J., DeGloria, S. D., & Elliot, R. (1999). A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences*, 5(1–4), 23–27.
- Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, 55(2), 95–107.
- Rossiter, D. G. (2008). Digital soil mapping as a component of data renewal for areas with sparse soil data infrastructures. In A. E. Hartemink, A. McBratney, & M. D. L. Mendonça-Santos (Eds), *Digital Soil Mapping with Limited Data* (pp. 69–80). Springer Netherlands. https://doi.org/10.1007/978-1-4020-8592-5_6
- Rossiter, D. G. (2018). Past, present & future of information technology in pedometrics. *Geoderma*, 324, 131–137. <https://doi.org/10.1016/j.geoderma.2018.03.009>
- Rossiter, D. G., & Poggio, L. (2025). Representing soil landscapes from digital soil mapping products – helping the map to speak for itself. *SOIL*, 11(2), 849–881. <https://doi.org/10.5194/soil-11-849-2025>
- Rossiter, D. G., Poggio, L., Beaudette, D., & Libohova, Z. (2022). How well does digital soil mapping represent soil geography? An investigation from the USA. *SOIL*, 8(2), 559–586. <https://doi.org/10.5194/soil-8-559-2022>
- Royal, A. (2023). Arid soils. In M. Brown, J. Burland, T. Chapman, K. Higgins, H. Skinner, & D. Toll (Eds), *ICE Manual of*

- Geotechnical Engineering: I* (2nd edn, pp. 343–372). Emerald Publishing Limited. <https://doi.org/10.1680/icemge.66816.0343>
- Rühlmann, J., & Körschens, M. (2009). Calculating the Effect of Soil Organic Matter Concentration on Soil Bulk Density. *Soil Science Society of America Journal*, 73(3), 876–885. <https://doi.org/10.2136/sssaj2007.0149>
- Safanelli, J., Poppiel, R., Ruiz, L., Bonfatti, B., Mello, F., Rizzo, R., & Demattê, J. (2020). Terrain analysis in Google Earth Engine: A method adapted for high-performance global-scale analysis. *ISPRS International Journal of Geo-Information*, 9(6), 400. <https://doi.org/10.3390/ijgi9060400>
- Samuel-Rosa, A., Heuvelink, G. B. M., Vasques, G. M., & Anjos, L. H. C. (2015). Do more detailed environmental covariates deliver more accurate soil maps? *Geoderma*, 243–244, 214–227. <https://doi.org/10.1016/j.geoderma.2014.12.017>
- Sayre, R., Dangermond, J., Frye, C., Vaughan, R., Aniello, P., Breyer, S. P., Cribbs, D., Hopkins, D., Nauman, R., Derrenbacher, W., Wright, D. J., Brown, C., Convis, C., Smith, J. H., Benson, L., VanSistine, D. P., Warner, H., Cress, J. J., Danielson, J. J., ... Comer, P. (2014). *A new map of global ecological land units—An ecophysiographic stratification approach*. Association of American Geographers; U. S. Geological Survey; GEO BON.
- Schaetzl, R. J., & Thompson, M. L. (2015). *Soils: Genesis and geomorphology* (2nd ed.). Cambridge University Press.
- Schillaci, C., Braun, A., & Kropáček, J. (2015). Terrain analysis and landform recognition. In *Geomorphological techniques*. British Society for Geomorphology.
- Schmidinger, J., & Heuvelink, G. B. M. (2023). Validation of uncertainty predictions in digital soil mapping. *Geoderma*, 437, 116585. <https://doi.org/10.1016/j.geoderma.2023.116585>
- Schmidinger, J., Vogel, S., Barkov, V., Pham, A.-D., Gebbers, R., Tavakoli, H., Correa, J., Tavares, T. R., Filippi, P., Jones, E. J., Lukas, V., Boenecke, E., Rühlmann, J., Schroeter, I., Kramer, E., Paetzold, S., Kodaira, M., Wadoux, A. M. J.-C., Bragazza, L., ... Atzmueller, M. (2025). LimeSoDa: A dataset collection for benchmarking of machine learning regressors in digital soil mapping. *Geoderma*, 459, 117337. <https://doi.org/10.1016/j.geoderma.2025.117337>
- Schoeneberger, P. J., Wysocki, D. A., Busskohl, C., & Libohova, Z. (2017). Landscapes, geomorphology, and site description. In C. Ditzler, K. Scheffe, & H. C. Monger (Eds), *Soil Survey Manual*. https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/ref/?cid=nrcs142p2_054252
- Schulze, R. E., & Horan, M. J. C. (2007). Altitude and relative relief. In R. E. Schulze (Ed.), *South African atlas of climatology and agrohydrology*. Water Research Commission.
- Shepherd, K. D., Ferguson, R., Hoover, D., van Egmond, F., Sanderman, J., & Ge, Y. (2022). A global soil spectral calibration library and estimation service. *Soil Security*, 7, 100061.
- Shrestha, D. L., & Solomatine, D. P. (2006). Machine learning approaches for estimation of prediction interval for the model output. *Neural Networks*, 19(2), 225–235. <https://doi.org/10.1016/j.neunet.2006.01.012>
- Simmonds, E. B. (1994). *Rural water supply maintenance in southern Kunene Region Namibia. Soils assessment*. [Technical Report]. InterConsult Namibia (Pty) Ltd. for Africare.
- Simmonds, E. B. (1998). *Soils of the Okavango river basin. Specialist report*. Interconsult, for the Permanent Okavango River Basin Commission (OKACOM).
- Simmonds, E. B., & Burke, A. (2000). *Soils of Kavango Region. Natural resource mapping of the Kavango Region. A report for the Kavango Environmental Profile*. [Technical Report]. Interconsult, for the Directorate of Environmental Affairs, Ministry of Environment and Tourism.
- Simmonds, E. B., & Forbes Irving, T. J. M. (1995). *Soils assessment and land evaluation. Rural water supply maintenance project in southern Kunene Region* [Technical Report]. Interconsult Namibia (Pty) Ltd for Africare.
- Simperegui, K. B. D., Kouame, A. K. K., Kwesie, B., Bindraban, P. S., Adzawla, W., Asamoah, E., & El Gharous, M. (2025). Digital mapping of Ghana's soil properties and nutrients: Performance of spline and weighted average approaches. *Geoderma*, 459, 117365. <https://doi.org/10.1016/j.geoderma.2025.117365>
- Soil Classification Working Group. (1991). *Soil classification: A taxonomic system for South Africa*. Department of Agricultural Development.
- Soil Survey Staff. (1999). *Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys* (2nd edn). Natural Resources Conservation Service, USDA.
- Soil Survey Staff. (2014). *Keys to soil taxonomy* (12th edn). Natural Resources Conservation Service, USDA.
- Soil Survey Staff. (2017). *Soil survey manual*. Government Printing Office. <https://www.nrcs.usda.gov/sites/default/files/2022-09/The-Soil-Survey-Manual.pdf>

- Stenberg, B., Viscarra Rossel, R. A., Mouazen, A. M., & Wetterlind, J. (2010). Visible and near infrared spectroscopy in soil science. In D. L. Sparks (Ed.), *Advances in Agronomy*. (pp. 163–215). Academic Press.
- Stoorvogel, J. J., Kempen, B., Heuvelink, G. B. M., & De Bruin, S. (2009). Implementation and evaluation of existing knowledge for digital soil mapping in Senegal. *Geoderma*, 149(1–2), 161–170. <https://doi.org/10.1016/j.geoderma.2008.11.039>
- Strohbach, B. J., Adank, W. F., Coetzee, M. E., & Jankowitz, W. J. (2019). A baseline description of the soils and vegetation of farm Klein Boesman, Khomas Region, Namibia. *Namibian Journal of Environment*, 3, 37–55.
- Strohbach, B. J., Coetzee, M., Marais, E., & Maggs-Kölling, G. (2025). *Ecosystems of Namibia map vs4.0* [Data set]. Zenodo. <https://doi.org/10.5281/ZENODO.17218337>
- Strohbach, B. J., & Kutuahupira, J. T. (2014). Vegetation of the eastern communal conservancies in Namibia: II. Environmental drivers. *Koedoe*, 56(1). <https://doi.org/10.4102/koedoe.v56i1.1117>
- Strohbach, B. J., Strohbach, M., Kutuahupira, J. T., & Mouton, H. D. (2004). *A reconnaissance survey of the landscapes, soils and vegetation of the eastern communal areas (Otjozondjupa and Omaheke Regions)*. National Botanical Research Institute, for the Desert Research Foundation of Namibia and the Desert Margins Programme.
- Strohbach, B. J., Strydom, E., Nesongano, C., Zimmermann, I., Cauwer, V., & Coetzee, M. (2023). *Assessment of the impact of bush encroachment and bush control on climate change mitigation and adaptation in Namibia. Final report*. AHT GROUP GmbH & Perivoli Rangeland Institute, for Deutsche Gesellschaft für Internationale Zusammenarbeit.
- Stum, A. K., Boettinger, J. L., White, M. A., & Ramsey, R. D. (2010). Random forests applied as a soil spatial predictive model in arid Utah. In J. L. Boettinger, D. W. Howell, A. C. Moore, A. E. Hartemink, & S. Kienast-Brown (Eds), *Digital soil mapping: Bridging research, environmental application, and operation* (pp. 179–190). Springer.
- Suleymanov, A., Gabbasova, I., Komissarov, M., Suleymanov, R., Garipov, T., Tuktarova, I., & Belan, L. (2023). Random forest modeling of soil properties in saline semi-arid areas. *Agriculture* 13(5), 976. <https://doi.org/10.3390/agriculture13050976>
- Tadono, T., Ishida, H., Oda, F., Naito, S., Minakawa, K., & Iwamoto, H. (2014). Precise global DEM generation by ALOS PRISM. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II–4, 71–76. <https://doi.org/10.5194/isprsannals-II-4-71-2014>
- Tang, C.-S., Zhu, C., Cheng, Q., Zeng, H., Xu, J.-J., Tian, B.-G., & Shi, B. (2021). Desiccation cracking of soils: A review of investigation approaches, underlying mechanisms, and influencing factors. *Earth-Science Reviews*, 216, 103586. <https://doi.org/10.1016/j.earscirev.2021.103586>
- Terhoeven-Urselmans, T., Vagen, G., Spaargaren, O., & Shepherd, K. D. (2010). Prediction of soil fertility properties from a globally distributed soil mid-infrared spectral library. *Soil Science Society of America Journal*, 74(5), 1792–1799. <https://doi.org/10.2136/sssaj2009.0218>
- Theobald, D. M., Harrison-Atlas, D., Monahan, W. B., & Albano, C. M. (2015). Ecologically-relevant maps of landforms and physiographic diversity for climate adaptation planning. *PloS One*, 10(12), e0143619.
- Thompson, J. A., Roecker, S., Grunwald, S., & Owens, P. R. (2012). Digital soil mapping: Interactions with and applications for hydrogeology. *Hydrogeology*, 1, 664–709.
- Titouine, M., & Meziane, T. (2015). The influence of altitude and landforms on some biochemical and hematological parameters in Ouled Djellal ewes from arid area of south east Algeria. *Veterinary World*, 8(1), 130–134. <https://doi.org/10.14202/vetworld.2015.130-134>
- Trabucco, A., & Zomer, R. (2022). *Global aridity index and potential evapotranspiration (ET0) database: Version 3* [Data set]. figshare. <https://doi.org/10.6084/M9.FIGSHARE.7504448.V5>
- Trippner, C. (1997). Salt content as an eco-pedological limiting factor in soils of the Etosha National Park, northern Namibia. *Madoqua*, 20(1), 105–113.
- Trippner, C. (1998). *Semi-detailed soil survey and landscape ecological risk evaluation in the south-western and central-western parts of the Etosha National Park / Namibia* (p. 659 4-1 4-2) [Project report].
- United Nations. (1992a). *United Nations Convention on Biological Diversity* (Issue ue 15). United Nations.
- United Nations. (1992b). *United Nations Framework Convention on Climate Change*. United Nations.
- United Nations. (1994). *United Nations Convention to Combat Desertification* (p. 58). United Nations.
- Van der Waal, C., Coetzee, M., Tjelos, N., Lubbe, L., & Knox, N. (2021a). Agro-ecological zone and carrying capacity maps—Final report. In *Project CS/RFQ/25-2/2017, for the Namibian Ministry of Agriculture, Water and Land Reform*. Agri-ecological Services.

- Van der Waal, C., Lubbe, L. K., Tjelos, N., & Coetzee, M. E. (2021b). *National potential carrying capacity map of Namibia* [Map]. Agri-ecological Services, for the Ministry of Agriculture, Water and Land Reform, Republic of Namibia. Project CS/RFQ/25-2/2017: Agro-Ecological Zone and Carrying Capacity Maps.
- Van der Waal, C., Lubbe, L., Knox, N., Tjelos, N., & Coetzee, M. E. (2021c). *Agro-ecological zones map of Namibia*. [Map]. Agri-ecological Services, for the Ministry of Agriculture, Water and Land Reform, Republic of Namibia. Project CS/RFQ/25-2/2017: Agro-Ecological Zone and Carrying Capacity Maps.
- Van Der Westhuizen, S., Heuvelink, G. B. M., Hofmeyr, D. P., & Poggio, L. (2022). Measurement error-filtered machine learning in digital soil mapping, *Spatial Statistics*, 47, 100572. <https://doi.org/10.1016/j.spasta.2021.100572>
- van Egmond, F., van der Woude, T., Turdukulov, U., van Genuchten, P., de Sousa, L., Kempen, B., Batjes, N. H., van Oostrum, A., Mantel, S., Kooiman, A., Ruiperez Gonzalez, M., Poggio, L., Genova, G., Bai, Z., & Toner, E. (with ISRIC - World Soil Information). (2023). *Development options for a soil information workflow and system*. ISRIC - World Soil Information. <https://doi.org/10.17027/ISRIC-TMKB-PR58>
- van Engelen, V. W. P., & Dijkshoorn, J. A. (Eds). (2013). *Global and national soils and terrain databases (SOTER). Procedures manual, version 2.0*. ISRIC–World Soil Information. http://www.isric.org/sites/default/files/isric_report_2013_04.pdf
- van Engelen, V. W. P., & Wen, T. T. (Eds). (1995). *Global and national soils and terrain digital databases—SOTER. Procedures manual* (World Soil Resources Report 74). UNEP; ISSS; ISRIC; FAO. <https://edepot.wur.nl/493802>
- Van Woerkom, J., & Louw, P. A. (1971). *N Ondersoek van grondprobleme op die Hardapnedersetting, S.W.A* (Verslag No. 754/14/71). NIGB.
- Vandana, N., Suresh, G. J. R., Mitran, T., & Mahadevappa, S. G. (2024). Digital mapping of soil pH and electrical conductivity using geostatistics and machine learning. *International Journal of Environment and Climate Change*, 14(2), 273–286. <https://doi.org/10.9734/ijccc/2024/v14i23944>
- Vaysse, K., & Lagacherie, P. (2017). Using quantile regression forest to estimate uncertainty of digital soil mapping products. *Geoderma*, 291, 55–64. <https://doi.org/10.1016/j.geoderma.2016.12.017>
- Verheye, W. (2022). Soils of arid and semi-arid areas. In *Land Use, Land Cover and Soil Sciences: VII*. UNESCO-EOLSS.
- Viscarra Rossel, R. A., Behrens, T., Ben-Dor, E., Brown, D. J., Demattê, J. A. M., Shepherd, K. D., Shi, Z., Stenberg, B., Stevens, A., Adamchuk, V., Aichi, H., Barthès, B. G., Bartholomeus, H. M., Bayer, A. D., Bernoux, M., Böttcher, K., Brodský, L., Du, C. W., Chappell, A., ... Ji, W. (2016). A global spectral library to characterize the world's soil. *Earth-Science Reviews*, 155, 198–230. <https://doi.org/10.1016/j.earscirev.2016.01.012>
- Vogel, H.-J., Bartke, S., Daedlow, K., Helming, K., Kögel-Knabner, I., Lang, B., Rabot, E., Russell, D., Stöbel, B., Weller, U., Wiesmeier, M., & Wollschläger, U. (2018). A systemic approach for modeling soil functions. *SOIL*, 4(1), 83–92. <https://doi.org/10.5194/soil-4-83-2018>
- Vogel, H.-J., Eberhardt, E., Franko, U., Lang, B., Ließ, M., Weller, U., Wiesmeier, M., & Wollschläger, U. (2019). Quantitative evaluation of soil functions: Potential and state. *Frontiers in Environmental Science*, 7, 164. <https://doi.org/10.3389/fenvs.2019.00164>
- Wadoux, A. M. J.-C., & Minasny, B. (2024). Some limitations of the concordance correlation coefficient to characterise model accuracy. *Ecological Informatics*, 83, 102820. <https://doi.org/10.1016/j.ecoinf.2024.102820>
- Wadoux, A. M. J.-C., Minasny, B., & McBratney, A. B. (2020). Machine learning for digital soil mapping: Applications, challenges and suggested solutions. *Earth-Science Reviews*, 210, 103359. <https://doi.org/10.1016/j.earscirev.2020.103359>
- Wadoux, A. M. J.-C., Odeh, I. O. A., & McBratney, A. B. (2023). Overview of pedometrics. In *Encyclopedia of Soils in the Environment* (pp. 471–485). Elsevier. <https://doi.org/10.1016/B978-0-12-822974-3.00001-X>
- Walker, P., Hall, G., & Protz, R. (1968). Relation between landform parameters and soil properties. *Soil Science Society of America Journal*, 32, 101–104.
- Walter, C., Lagacherie, P., & Follain, S. (2006). Integrating pedological knowledge into digital soil mapping. In *Developments in Soil Science* (Vol. 31, pp. 281–615). Elsevier. [https://doi.org/10.1016/S0166-2481\(06\)31022-7](https://doi.org/10.1016/S0166-2481(06)31022-7)
- Wang, Q., Moreno-Martínez, Á., Muñoz-Marí, J., Campos-Taberner, M., & Camps-Valls, G. (2023). Estimation of vegetation traits with kernel NDVI. *ISPRS Journal of Photogrammetry and Remote Sensing*, 195, 408–417. <https://doi.org/10.1016/j.isprsjprs.2022.12.019>
- Webster, R. (1994). The development of pedometrics. *Geoderma*, 62(1–3), 1–15. [https://doi.org/10.1016/0016-7061\(94\)90024-8](https://doi.org/10.1016/0016-7061(94)90024-8)
- Webster, R. (2000). Is soil variation random? *Geoderma*, 97(3–4), 149–163. [https://doi.org/10.1016/S0016-7061\(00\)00036-7](https://doi.org/10.1016/S0016-7061(00)00036-7)

- Weiss, A. (2001). *Topographic position and landforms analysis*. ESRI User Conference, San Diego, CA.
- Wiesmeier, M., Barthold, F., Blank, B., & Kögel-Knabner, I. (2011). Digital mapping of soil organic matter stocks using random forest modeling in a semi-arid steppe ecosystem. *Plant Soil*, 340, 7–24.
- Wild, S. (2016). Quest to map Africa's soil microbiome begins. *Nature*, 539(7628), 152–152. <https://doi.org/10.1038/539152a>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>
- Wilson, M. F. J., O'Connell, B., Brown, C., Guinan, J. C., & Grehan, A. J. (2007). Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, 30, 3–35.
- Wösten, J. H. M. (1997). Pedotransfer functions to evaluate soil quality. In *Developments in Soil Science* (Vol. 25, pp. 221–245). Elsevier. [https://doi.org/10.1016/S0166-2481\(97\)80037-2](https://doi.org/10.1016/S0166-2481(97)80037-2)
- Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., & Pavelsky, T. M. (2019). MERIT hydro: A high-resolution global hydrography map based on latest topography dataset. *Water Resources Research*, 55(6), 5053–5073. <https://doi.org/10.1029/2019WR024873>
- Yigini, Y., Olmedo, G. F., Reiter, S., Baritz, R., Viatkin, K., & Vargas, R. (Eds). (2018). *Soil organic carbon mapping cookbook* (2nd edn, p. 220). Food and Agriculture Organization of the United Nations.
- Zádorová, T., Skála, J., Žížala, D., Vaněk, A., & Penížek, V. (2021). Harmonization of a large-scale national soil database with the world reference base for soil resources 2014. *Geoderma*, 384, 114819. <https://doi.org/10.1016/j.geoderma.2020.114819>
- Zádorová, T., Penížek, V., Mihaljevič, M., Koubová, M., Lisá, L., Ettler, V., Tejnecký, V., Drábek, O., Pavlů, L., Křibek, B., Vaněk, A., Sracek, O., Rojas, J. R., Hrdlička, T., Vokurková, P., & Mapani, B. (2025). Local diversity of soil forming processes in the semi-arid tropics and its environmental drivers: An example from Otavi Mountains, northern Namibia. *CATENA*, 249, 108671. <https://doi.org/10.1016/j.catena.2024.108671>
- Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsendbazar, N.-E., Xu, P., Ramoino, F., & Arino, O. (2022). *ESA WorldCover 10 m 2021 v200* (Version v200) [Data set]. Zenodo. <https://doi.org/10.5281/ZENODO.7254221>
- Zhang, G., Liu, F., & Song, X. (2017). Recent progress and future prospect of digital soil mapping: A review. *Journal of Integrative Agriculture*, 16(12), 2871–2885. [https://doi.org/10.1016/S2095-3119\(17\)61762-3](https://doi.org/10.1016/S2095-3119(17)61762-3)
- Zomer, R. J., Xu, J., & Trabucco, A. (2022). Version 3 of the global aridity index and potential evapotranspiration database. *Scientific Data*, 9(1), 409. <https://doi.org/10.1038/s41597-022-01493-1>

B. Environmental covariates assembled in the predictor stack

Covariate name in scripts	Description	Dataset Name; Source; Provider	Reference	Original resolution (m)
dem	NASADEM Merged Global 1 arc second Digital Elevation Model V001; unit: metres. NASADEM_HGT is based on the 2000 Shuttle Radar Topography Mission (SRTM), reprocessed with auxiliary data from ASTER GDEM, ICESat GLAS, and ALOS PRISM, to extend spatial coverage over polar regions, improve surface elevation measurements and geolocation accuracy, fill voids, and convert it to the EGM96 geoid.	NASADEM HGT v001; NASA /USGS/JPL-Caltech; Google Earth Engine Catalogue	(NASA JPL, 2020). https://doi.org/10.5067/MEaSURES/NASADEM/NASADEM_HGT.001 ; https://developers.google.com/earth-engine/datasets/catalog/NASA_NASADEM_HGT_001 ; https://lpdaac.usgs.gov/products/nasadem_hgtv001/	30.92
Slope	Gradient or rate of change of elevation between neighbouring cells; unit: degrees	Derived from NASADEM in Google Earth Engine, using TAGEE script	(Safanelli et al., 2020). https://doi.org/10.3390/ijgi9060400 ; https://github.com/zecojls/tagee	30.92
Aspect	Represents exposure faces; unit: degrees (0–360)			
Eastness	sin(aspect); emphasises the continuous east-west gradient of the slope			
Northness	cos(aspect); emphasises the continuous north–south gradient of the slope			
HorizontalCurvature	Plan curvature; measured perpendicular to the slope direction; controls lateral flow divergence/convergence; concave (-): flow converges, causing erosion channels and deposition zones; convex (+): flow diverges, reducing erosion, promoting sheet flow			
VerticalCurvature	Profile curvature; measured along the slope direction; concave (-): enhances water accumulation and slows down flow; convex (+): increases flow acceleration and enhances erosion.			
chili	JAXA Global ALOS Continuous Heat-Insolation Load Index (CHILI); surrogate for effects of insolation and topographic shading on evapotranspiration	JAXA; Global ALOS CHILI; Global ALOS mTPI; Global ALOS Topographic Diversity; Global ALOS Landforms; Conservation Science Partners; Google Earth Engine	(Theobald et al., 2015). https://developers.google.com/earth-engine/datasets/catalog/CSP_ERGo_1_0_Global_ALOS_CHILI ; https://developers.google.com/earth-engine/datasets/catalog/CSP_ERGo_1_0_Global_ALOS_mTPI ; https://developers.google.com/earth-engine/datasets/catalog/CSP_ERGo_1_0_Global_ALOS	90
tpi	JAXA Global ALOS Multi-Scale Topographic Position Index (mTPI); distinguishes ridges (+) from valleys (-)			270
topo_diver	JAXA Global ALOS Topographic Diversity; represents the variety of temperature and moisture conditions available to species as local habitat			270
landforms_alos	JAXA Global ALOS Landforms; hillslope positions, as			90

	influenced by insolation and/or shading in relation to impact on ecological processes; 15 classes	Catalogue	topoDiversity; https://developers.google.com/earth-engine/datasets/catalog/CSP_ERGo_1_0_Global_ALOS_landforms .	
flow_dir	Flow direction / local drainage direction; band 'dir'	MERIT Hydro: Global Hydrography Dataset; Dai Yamazaki (University of Tokyo); Google Earth Engine Catalogue	(Yamazaki et al., 2019). https://doi.org/10.1029/2019WR024873 . https://developers.google.com/earth-engine/datasets/catalog/MERIT_Hydro_v1_0_1/	92.77
hand	Height above the nearest drainage (m) / hydrologically adjusted elevations above EGM96 geoid; 10 cm vertical increments); band 'hand'			
flow_accumul	Upstream drainage area (flow accumulation area); unit: km ² ; band 'upa'			92.77
river_dist	Distance to nearest drainage line; implications for connectivity and flow dynamics	HYDROSHEDS (Hydrological data and maps based on Shuttle Elevation Derivatives at multiple Scales) Core Data v1.1; WWF US	(Lehner et al., 2008) https://doi.org/10.1029/2008eo100001 . https://www.hydrosheds.org/hydrosheds-core-downloads	463.83
flow_lend_d	Flow length downstream: distance (m) from the final pour point of the river in downstream direction, at the ocean or an inland sink (LDN); 15" resolution			
flow_len_up	Flow length upstream: distance (m) from the furthest river source in upstream direction (LUP); 15" resolution			
landcover	Global land cover map (2021); 10 m resolution; based on Sentinel-1 and Sentinel-2 data; 11 land cover classes	ESA WorldCover 10m 2021 v200; ESA/VITO/Brockmann Consult /CS/GAMMA Remote Sensing/ IIASA/WUR; Google Earth Engine Catalogue	© ESA WorldCover project / Contains modified Copernicus Sentinel data (2021) processed by ESA WorldCover consortium. (Zanaga et al., 2022). https://doi.org/10.5281/zenodo.7254221 ; https://developers.google.com/earth-engine/datasets/catalog/ESA_WorldCover_v200	9.28
Prec_wc2	WorldClim Climatology V2; 1970–2000; mean annual precipitation; unit: mm	WorldClim Climatology V2	https://www.worldclim.org/data/worldclim21.html (Fick & Hijmans, 2017). https://doi.org/10.1002/joc.5086 .	927.67
tavg_wc2	WorldClim Climatology V2; 1970–2000; mean annual temperature; mean of average monthly temperature; unit: °C			927.67
arid_ind	Aridity Index; ratio of precipitation to potential evapotranspiration (FAO-56 Penman-Monteith Reference Evapotranspiration ET0); aggregated annual values 1970–2000; correction factor 0.0001	Global Aridity Index and Potential Evapo-Transpiration (ET0) Database Version 3 (Global-AI_PET_v3); CGIAR-CSI	https://doi.org/10.6084/m9.figshare.7504448.v5 https://csidotinfo.wordpress.com/data/global-aridity-and-pet-database/ https://www.nature.com/articles/s41597-022-01493-1	927.67
pet	Potential evapotranspiration (1970–2000) based on FAO-56 Penman-Monteith Reference Evapotranspiration (ET0) equation			
Landsat surface	Landsat 5 surface reflectance (1985-01-01 - 2011-12-31);	USGS Landsat 5 Level	Landsat-5 images courtesy of the U.S. Geological	30

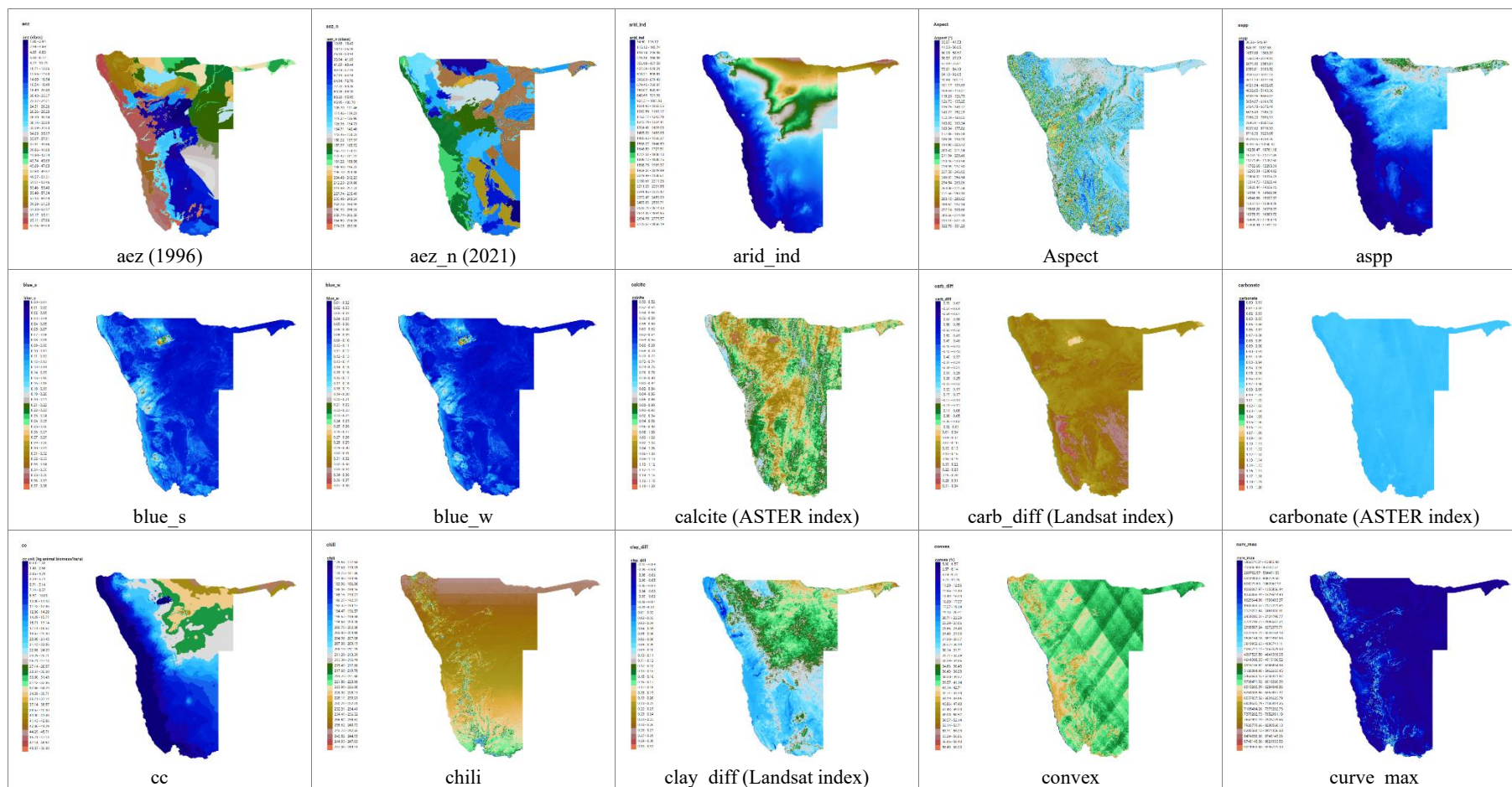
reflectance: blue_s blue_w green_s green_w red_s red_w nir_s nir_w swir1_s swir1_w swir2_s swir2_w (_s: summer; January and February) (_w: winter; June and July)	SR_B1 (0.45–0.52 μm, B) SR_B2 (0.52–0.60 μm, G) SR_B3 (0.63–0.69 μm, R) SR_B4 (0.77–0.90 μm, NIR) SR_B5 (1.55–1.75 μm, SWIR1) SR_B7 (2.08–2.35 μm, SWIR2)	2, Collection 2, Tier 1; USGS Landsat 8 Level 2, Collection 2, Tier 1; United States Geological Survey; Google Earth Engine Catalogue	Survey, https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LT05_C02_T1_L2	
	Landsat 8 surface reflectance, Level 2, Collection 2, Tier 1 (2013-04-01, 2024-07-25): SR_B2 (0.452–0.512 μm, B), SR_B3 (0.533–0.590 μm, G), SR_B4 (0.636–0.673 μm, R), SR_B5 (0.851–0.879 μm, NIR), SR_B6 (1.566–1.651 μm, SWIR1), SR_B7 (2.107–2.294 μm, SWIR2),		Landsat-8 images courtesy of the U.S. Geological Survey. https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C02_T1_L2	30
ndvi_s ndvi_w savi_s savi_w msavi_s msavi_w evi_s evi_w kndvi_s kndvi_w	Normalized Difference Vegetation Index; NDVI = ((NIR - R) / (NIR + R)) Soil Adjusted Vegetation Index; SAVI = ((NIR - R) / (NIR + R + L)) * (1 + L), L is soil brightness correction factor; here defined as 0.5 Modified Soil Adjusted Vegetation Index; MSAVI = (2 * NIR + 1 - sqrt((2 * NIR + 1) ² - 8 * (NIR - R))) / 2 Enhanced Vegetation Index; EVI = 2.5 * ((NIR - R) / (NIR + 6 * R - 7.5 * B + 1)) Kernel Normalized Difference Vegetation Index; kNDVI = tanh(ndvi ²) = tanh(((NIR - R) / (NIR + R)) ²)	Vegetation indices derived from Landsat 5 and 8 in Google Earth Engine	NDVI: https://www.usgs.gov/landsat-missions/landsat-normalized-difference-vegetation-index SAVI: https://www.usgs.gov/landsat-missions/landsat-soil-adjusted-vegetation-index . MSAVI: https://www.usgs.gov/landsat-missions/landsat-modified-soil-adjusted-vegetation-index . EVI: https://www.usgs.gov/landsat-missions/landsat-enhanced-vegetation-index . kNDVI: https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-2/kndvi/ (Camps-Valls et al., 2021; Wang et al., 2023). https://github.com/IPL-UV/kNDVI?tab=readme-ov-file .	30
carb_diff clay_diff ferr_diff iron rock_out	Carbonate Normalization (or Difference) Ratio; (R - G) / (R + G) Clay Normalization Ratio; (SWIR1 - SWIR2) / (SWIR1 + SWIR2) Ferrous Minerals Normalization Ratio; (SWIR1 - NIR) / (SWIR1 + NIR) Iron Normalization Ratio; (R - SWIR2) / (R + SWIR2) Rock Outcrop Normalization Ratio; (SWIR1 - G) / (SWIR1 + G)	Mineralogy indices derived from Landsat 5 and 8 in Google Earth Engine	(Bodily, 2005; Boettinger et al., 2008; Jensen, 2005; Kienast-Brown et al., 2017; Nield et al., 2007; Stum et al., 2010)	

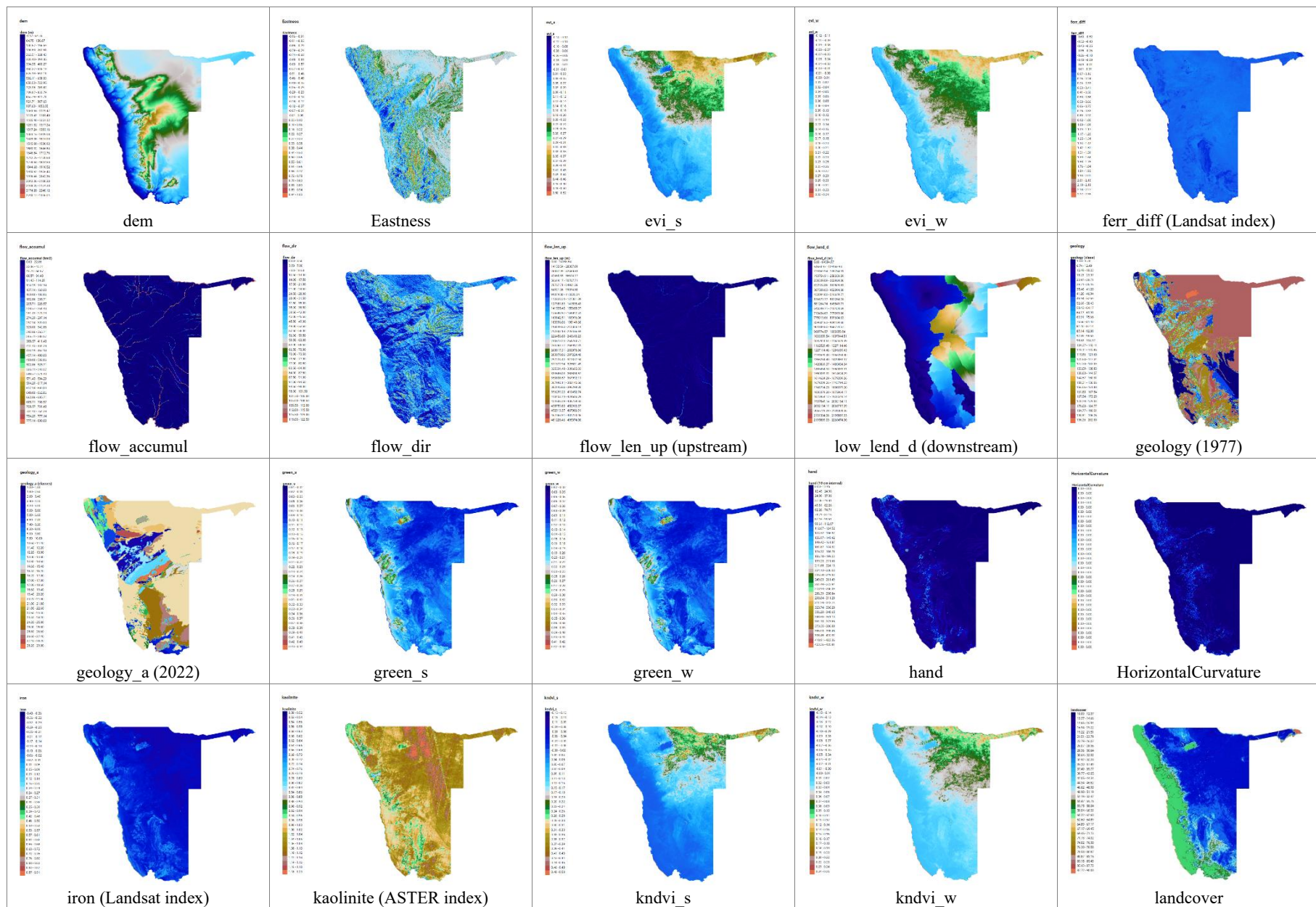
ASTER surface reflectance, mineralogical indices :		Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Level 1 Precision Terrain Corrected Registered At-Sensor Radiance (AST_L1T); USGS; Google Earth Engine Datasets; Indices derived from ASTER in Google Earth Engine	https://lpdaac.usgs.gov/products/ast_11tv003/ ; https://terra.nasa.gov/data/aster-data Algorithm by (Ahrari, 2024) (https://github.com/AmirhosseinAhrari/GoogleEarthEngine/blob/main/00067_aster_lithology), following (Rajendran & Nasir, 2019)	90
kaolinite index	(B4/B5)(B8/B6)			
calcite index	(B6/B8)(B9/B8)			
quartz index	(B11*B11)(B10*B12)			
carbonate index	(B13/B14)			
mafic index	(B12/B13)			
	Where B4 (1.6–1.7 µm) B5 (2.145–2.185 µm) B6 (2.185–2.225 µm) B8 (2.295–2.365 µm) B9 (2.36–2.43 µm) B10 (8.125–8.475 µm) B11 (8.475–8.825 µm) B12 (8.925–9.275 µm) B13 (10.25–10.95 µm) B14 (10.95–11.65 µm)			
Aez	Agro-ecological zones of Namibia, 1996, Categorical map, rasterised; based on land systems and growing period zones	AEZ (1996); Ministry of Agriculture, Water and Rural Development	(De Pauw, 1996; De Pauw et al., 1999a, 1999b; De Pauw & Coetzee, 1999)	93.78
aez_n	New Agro-ecological zones of Namibia, 2021; probabilistic raster map of land suitability for growing particular crops	AEZ (2021); MAWLR AEZ/CC Project CS/RFQ/25-2/2017: Agro-Ecological Zone and Carrying Capacity Maps	(Van der Waal et al., 2021a, 2021c)	231.92
cc	Potential carrying capacity map of Namibia, 2021		(Van der Waal et al., 2021b).	111.32
namsoil_13	National soil map of Namibia; 13 classes: reference soil group of dominant soil, World Reference Base 2015		(current study)	93.78
aspp	Average seasonal plant productivity 1999/2000–2018/19, based on Copernicus Global Land Operations Dry Matter Productivity	Atlas of Namibia, 2022	(Atlas of Namibia Team, 2022) aspp: http://the-eis.com/elibrary/search/26573 veg_types: http://the-eis.com/elibrary/search/26569 geology: http://the-eis.com/elibrary/search/26504	993.9
veg_types	Vegetation types			93.78
geology_a	Major groups of rocks, each group being of the same type formed at similar times and by similar processes			0.97
geology	Lithology units, from digital compilation of Geological Map of Namibia	Geological Survey of Namibia	(Geological Survey of Namibia, n.d.)	93.78
Landform_iwa	Iwahashi-Pike topographic classification (landform);	Iwahashi-Pike Global	https://esdac.jrc.ec.europa.eu/content/global-landform-	932.43

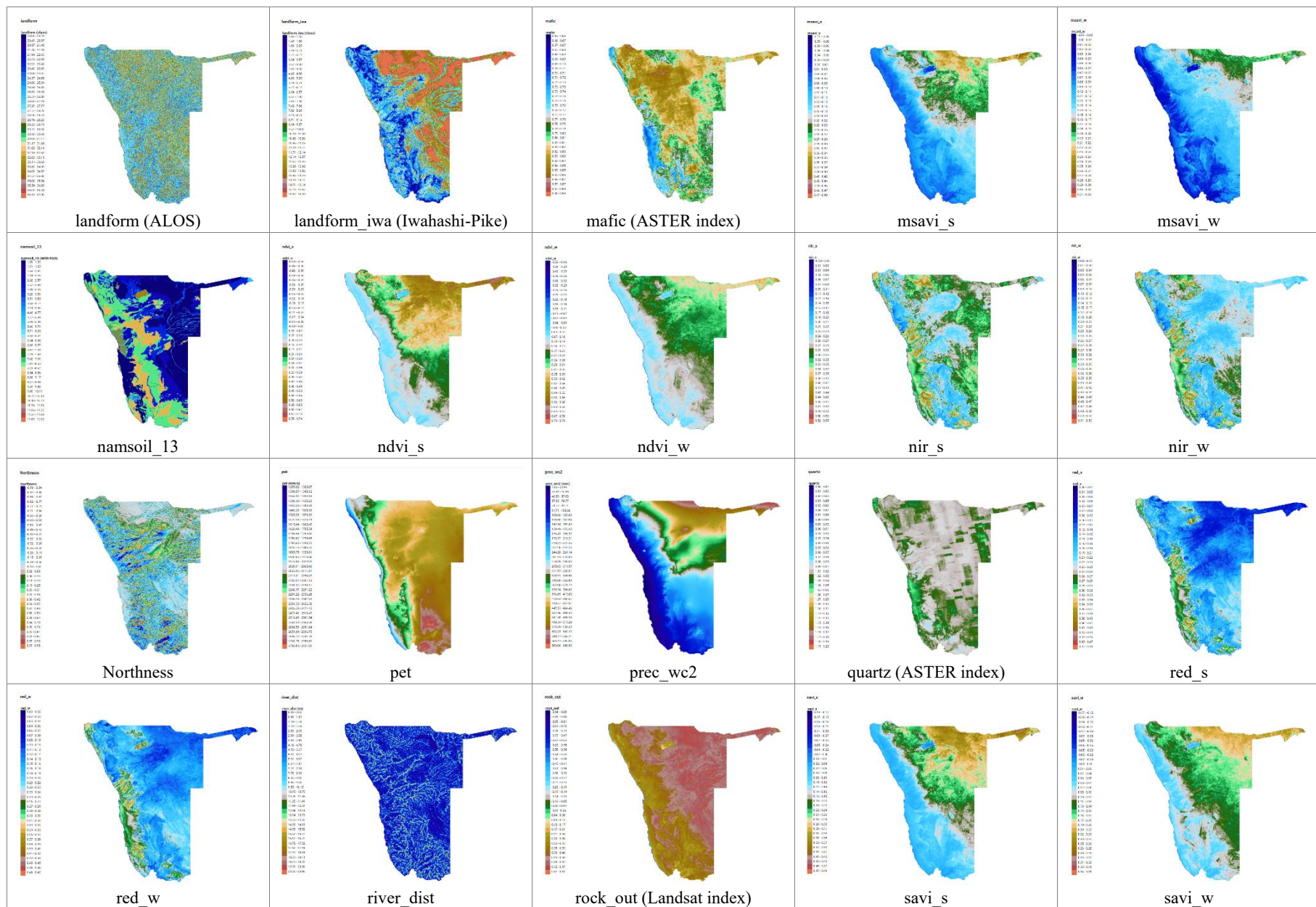
	relief classes classified with an unsupervised nested-means algorithms and a three part geometric signature, using slope gradient, surface texture, and local convexity on the SRTM30 DEM	Landform Classification; European Soil Data Centre (ESDAC)	classification (Iwahashi et al., 2018; Iwahashi & Pike, 2007)	
convex	Terrain convexity; ratio of number of cells with positive curvature to number of all valid cells within specified search radius; influences flow speed and convergence or divergence, erosivity	Derived in GIS from SRTM30 DEM		29.53
curv_max	Terrain curvature; 2nd derivative of slope; rate of change in slope; describes the convexity or concavity of the land surface	Derived in GIS from SRTM30 DEM		92.41

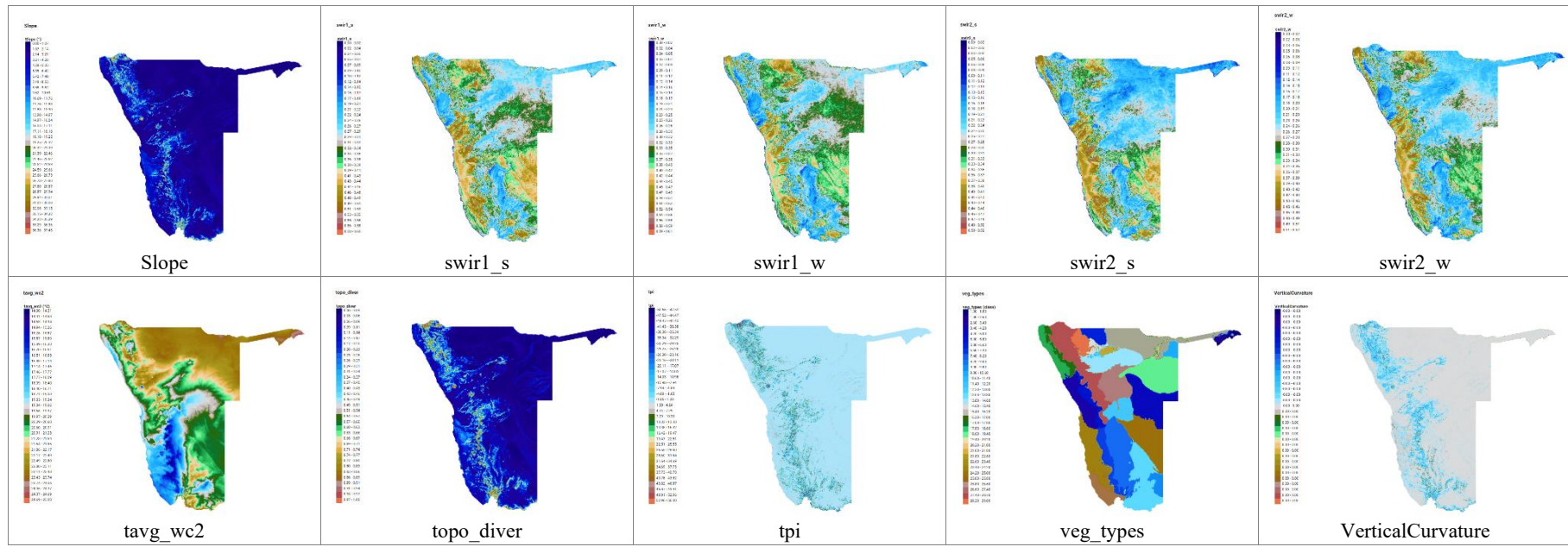
C. Thumbnail maps of all environmental predictors.

Names are as they occur in scripts. In case of ambiguity, more information is provided in brackets.









D. Digital soil mapping pipeline

Pipeline structure

Module	Folder	Description
1	1_data_treatment/	Soil profile data cleaning, table joining, missing-value treatment, and equal-area spline depth harmonisation to the three standard depth intervals. Includes exploratory data assessment scripts, histograms, boxplots, and spatial distribution plots.
2	2_feature_selection_tuning/	Boruta-based feature selection and Random Forest hyperparameter tuning (ntree, mtry, nodesize, sampsize) using the regression matrix exported from GEE. Outputs per-property configuration files consumed by the GEE modelling step.
2.5	2.5_GEE/	Google Earth Engine scripts for covariate preparation (lithological indices, Landsat spectral bands, ASTER mineral indices, terrain and climate layers), regression matrix export, and the core Random Forest modelling with 20-iteration bootstrap. Outputs Cloud-Optimised GeoTIFFs (mean, 5 th and 95 th percentiles) and per-iteration performance metrics. The full GEE repository is accessible at: https://code.earthengine.google.com/?accept_repo=users/Namibia_map/Soil_properties or via <code>git clone https://earthengine.googleusercontent.com/users/Namibia_map/Soil_properties</code> .
3	3_final_maps/	Post-processing of GEE outputs: applies physically meaningful min/max limits per property, computes the 90% prediction interval width (PI ₉₀), and generates publication-ready cartographic plots (PNG).
4	4_PICP/	Computes the 90% Prediction Interval Coverage Probability (PICP ₉₀) to evaluate whether the bootstrap-derived uncertainty bands contain the observed values at the expected nominal rate.
5	5_modeling_spectra_data_sfa/	Independent spectral modelling workflow: merges mid-infrared (MIR) spectra with wet chemistry reference data from the Soils for Africa (S4A) project, trains predictive models (Cubist, PLSR, Random Forest), and summarises results. Used to generate external validation predictions.
6	6_external_validation_sfa/	External validation: extracts predicted values from the Namibia maps, SoilGrids, and iSDAsoil at independent sample locations; harmonises units and depths; computes validation metrics (RMSE, R ² , CCC) and produces comparison plots.

Repository File Tree

```
Namibia_soil_maps_2024_2026/
├── Project_Overview_pipeline_environment.txt
├── 1_data_treatment/
│   ├── input/
│   │   ├── NamProfCleaned_21Mar2024.xlsx
│   │   └── National_boundary.shp (.dbf, .prj, .shx)
│   ├── metadata/
│   │   └── selected_soil_properties.txt
│   ├── output/
│   │   ├── datasets/ (splines, exploratory, GEE-ready exports)
│   │   ├── plots/    (histograms, boxplots, spatial distributions)
│   │   └── shapes/   (soil point shapefiles per property)
│   └── script/
│       ├── 1_Soil_data_join_tables_v2_preparation_for_splines.R
│       ├── 2_Soil_data_treatment_missing_values_depth_function_splines_new_0_30_60_100.R
│       ├── 2.1_Plots_centroids_Soil_data_treatment_..._splines_0_5_15_30_60_100_200_.R
│       ├── 3_plot_depth_hist.R
│       ├── 4_plot_hist_boxplot_splined_data_0_30_60_100.R
│       └── Exploratory_analyses_plots/
│           ├── Soil_data_assessment_..._Soil_classes_complete_info_OC_pH_BS_Sand_Clay_EC_BD.R
│           ├── Soil_data_assessment_..._Subsurface.R
│           └── Soil_data_assessment_..._Top_horiz.R
├── 2_feature_selection_tuning/
│   ├── input/
│   │   └── extracted_points_Nam_0_30_60_100_cm.csv
│   ├── output/
│   │   ├── plots/
│   │   └── properties_settings/ (per-property tuning configs)
│   └── script/
│       ├── 1_tuning_feature_selection_rf.R
│       ├── 2_check_data.R
│       └── 3_rename_github_gee_js.R
```

- └─ 2.5_GEE/
 - └─ input/ → GEE assets (see Readme.txt)
 - └─ outputs/ → Google Drive / GEE assets
 - └─ script/
 - └─ 1_lithological_indexes
 - └─ 2_gee_covariates_preparation
 - └─ 3_Export_regression_matrix_to_R_tunning
 - └─ 4_Extract_modeling_metrics_COG
 - └─ 5_thumbnail_all_covars_legend (independent/optional)
 - └─ properties_settings/
 - └─ as_txt/ (48 config files, one per property x depth)
 - └─ as_js_for_GEE/
- └─ 3_final_maps/
 - └─ input/
 - └─ earth_engine_Nam_modeling_output_*/ (GeoTIFFs + metrics CSVs)
 - └─ output/
 - └─ dataset/ (prop.limits.csv, ref_vals.csv, summary.limits.csv)
 - └─ plots/ (PNG maps: mean, lower, upper, range)
 - └─ script/
 - └─ make_maps_nam_v7_min_max_lg.R
- └─ 4_PICP/
 - └─ input/
 - └─ output/plots/
 - └─ script/
 - └─ 0_PICP_source_function.R
 - └─ 1_PICP.R

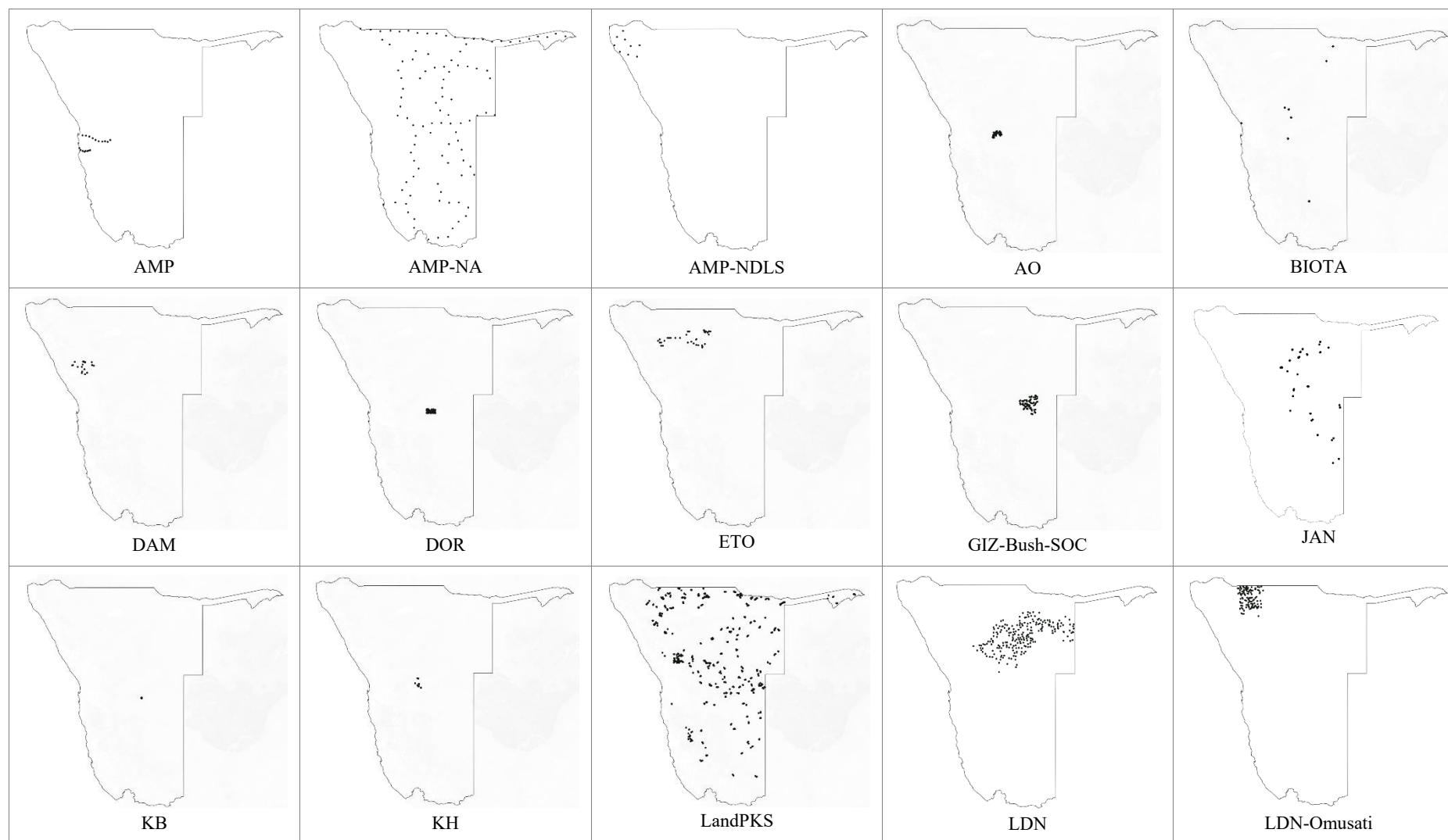
```

.
├── 5_modeling_spectra_data_sfa/
│   ├── input/
│   │   ├── s4a_submissions_Namibia_2023-09-28.xlsx
│   │   ├── S4A_Wetchem_20240911_NA.xlsx
│   │   └── spectral_data_Namibia_20250716.csv
│   ├── output/
│   │   ├── dataset/
│   │   ├── graphs_plots/
│   │   ├── metadata/
│   │   └── modeling_output/
│   └── script/
│       ├── 0_plot_fun_source.R
│       ├── 1_merge_spectra_mir_wetchem.R
│       ├── 2_modeling.R
│       └── 3_summarize_results.R
└── 6_external_validation_sfa/
    ├── input/
    │   ├── s4a_submissions_Namibia_2023-06-23.csv
    │   ├── S4A_Wetchem_20240911_NA.xlsx
    │   ├── S4A_Wetchem_20240911_NA_Topsoil_Samples_Coordinates.csv
    │   └── units_conversion.csv
    ├── output/plots/
    └── script/
        ├── 1_extract_samples_points__soil_grids_isda_Nam_maps.R
        ├── 2_units_conversion_depths_interpolation_v3_loop_datasets.R
        └── 3_validation_metrics_plots.R

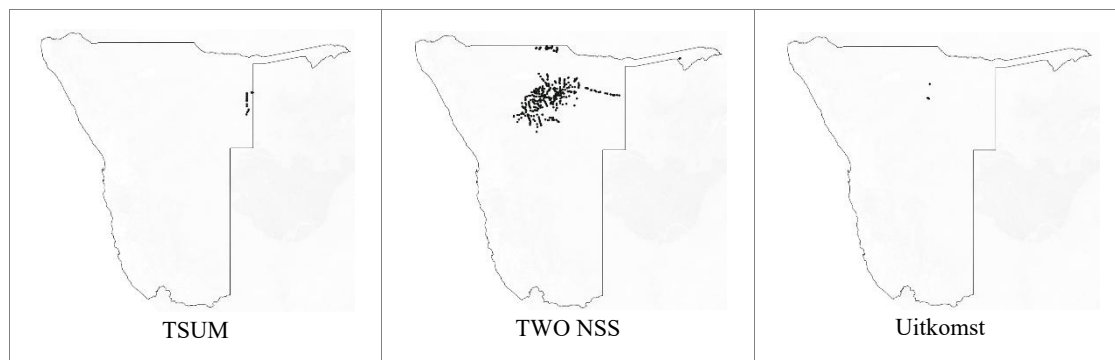
```

Source: Gelsleichter & Coetzee, 2026

E. Distribution of the NSPD's soil data points per dataset.







F. NSPD MS Access tables and fields.

Metadata

Dataset	CRS
Number	Detailed Description
Date	Reference
Type	Hyperlink
Survey or Project	Literature
Collector(s)/Author(s)	Original Data
Owner	Notes

Profile Registration

PRID	Day	Degree Square
Dataset	Country ID	Quarter Degree Square
Latitude	Data Origin	Farm No
Longitude	Surveyor(s)	Farm Name
Coordinate Reference System	Status	Location
Elevation m GPS	Profile Classification Original	Notes
Elevation m SRTM	Profile Classification Original Code	PRID Let
Date (YYYY-MM-DD)	Classification System	PRID Dig
Date (DD-Mon-YY)	WRB2015 Reclassification RSG	Alt PRID
Year	WRB2015 Reclassification Principal Qualifiers	Photos
Month	WRB2015 Reclassification Supplementary Qualifiers	

Site Description

PRID	Grazer Type	Surface Sealing Thickness Desc
Latitude	Human Influence Code	Surface Cracks Width Code
Longitude	Human Influence Desc	Surface Cracks Width cm
Diagnostic Horizon_Property_Material	Rock Outcrops Abundance_Cover Code	Distance Between Surface Cracks m
Present Weather Code	Rock Outcrops Cover Perc	Surface Salinity
Present Weather Desc	Rock Outcrops Cover Dec	Bleached Sand Code
Past Weather Code	Rock Outcrops m between	Bleached Sand Perc
Past Weather Desc	Rock Outcrops Type	Bleached Sand Desc
Soil Temperature Regime Code	Coarse Surface Fragments Cover Code	Flooded longer than 2 weeks per year
Soil Temperature Regime Desc	Coarse Surface Fragments Cover Perc	Major Landform Code 2nd level
Soil Moisture Regime Code	Coarse Surface Fragments Cover Desc	Major Landform Desc
Soil Moisture Regime Desc	Coarse Surface Fragments Size Code	Minor Landform Desc
Slpe Gradient Class	Coarse Surface Fragments Size cm	Position Code
Slope Form Vert	Coarse Surface Fragments Size Desc	Position Desc
Slope Form Hor	Coarse Surface Fragments Type	Vegetation
Slpe Gradient Measured Perc	Erosion Type Code	Land Cover

Slope Gradient Desc	Erosion Type	Parent Material
Gradient Perc from SRTM for Landform	Erosion Area Affected Code	Lithology
Slope Perc SRTM	Erosion Area Affected Perc	Drainage Code
Relief m_km	Erosion Degree Code	Drainage Desc
Drainage Dansity	Erosion Degree Desc	Geology
Present Land Use Code	Erosion Status Code	Effective Depth
Present Land Use Desc	Erosion Status Desc	Depth Class
Past Land Use Code	Surface Sealing Thickness Code	Biological Activity
Past Land Use Desc	Surface Sealing Thickness mm	Notes

Namibian Soil Literature

Author
Date
Title
Notes
Attachments

Environment

PRID	NSS1M Epigraph	OWA Soil1	Landscapes Atlas 2000
Latitude	NSS Land Division	OWA Soil2	Landscapes Short Atlas 2000
Longitude	NSS Land Province	OWA Soil3	Land Surface Forms USGS MORAP
AEZ 1998 Code	NSS Land Region	OWA Soil4	Landforms Iwahashi Pike
AEZ 1998 Desc	NSS1M Soil Unit	OWA Slope Description	Landforms Meybeck
GPZ 1998 Code	NSS1M Soil 1	OWA Texture	GEO Lithology Code
GPZ 1998 Desc	NSS1M Soil 2	OWA Surface Fragments	GEO Main Lithology
Catchment Strohbach	NSS1M Soil 3	OWA Salinity	GEO Age
Final Catchment Strohbach	NSS1M Soil 4	OWA Sodidity	GEO Other Rocks
Vegetation Type Giess	NSS1M Soil 5	OKA Epigraph	GEO Supergroup
CC 2003 kgAB_ha_a	NSS1M Soil 6	OKA LandUnit1	GEO Group
CC 2021 kgAB_ha_a	NSS1M Soil 7	OKA LandUnit 2	GEO Subgroup
FHFA Code	NSS1M Dissection	OKA SoilTypes	GEO Formation
FHFA Desc	NSS1M Relative Relief	OKA Soil1	GEO Complex
FHFA Potential	OWA Epigraph	OKA Soil2	GEO Suite
Slope perc SRTM	OWA LandUnit1	OKA Slope Description	GEO Intrusive
Ruggedness SRTM	OWA LandUnit 2	OKA Texture	GEO Unit
Aspect Deg SRTM	OWA LandUnit 3	OKA Salinity	GEO Lithology Class
Soil Unit 2021 Coetzee	OWA SoilTypes	OKA Sodidity	GEO Metamorphism

Horizons Field Data

PRID	Coarse Fragments Abundance Desc	Consistence Dry Code	Cementation_Compaction Continuity Code
HONU	Coarse Fragments Abundance Perc	Consistence Dry Desc	Cementation_Compaction Continuity Desc
Latitude	Coarse Fragments Size Code	Consistence Moist Code	Cementation_Compaction Fabric Code
Longitude	Coarse Fragments Size Desc	Consistence Moist Desc	Cementation_Compaction Fabric Desc
Representative Profile	Coarse Fragments Size mm	Consistence Wet Stickiness Code	Mineral Concentrations Kind Code
Diagnostic Horizon	Coarse Fragments Shape Code	Consistence Wet Stickiness Desc	Mineral Concentrations Kind Desc
Diagnostic Property	Coarse Fragments Shape Desc	Consistence Wet Plasticity Code	Mineral Concentrations Nature Code
Diagnostic Material	Coarse Fragments Degree of Weathering	Consistence Wet Plasticity Desc	Mineral Concentrations Nature Desc
Qualifier	Coarse Fragments Type	Porosity Measured	Mineral Concentrations Abundance Code
Horizon Code Short	Drainage Code	Porosity Class Code	Mineral Concentrations Abundance Desc
Horizon Code Original	Drainage Desc	Porosity Class Desc	Mineral Concentrations Abundance Perc
Subordinate Horizon Code Original	Field Texture Code	Porosity Class Perc	Mineral Concentrations Size Code
Lithology of R_C	Field Texture Desc	Voids Type Code	Mineral Concentrations Size Desc
Horizon Thickness cm	Carbonate Reaction Code	Voids Type Desc	Mineral Concentrations Shape Code
Upper Depth cm	Carbonate Reaction Desc	Voids Size Class	Mineral Concentrations Shape Desc
Lower Depth cm	Carbonate Reaction Perc	Voids Size Desc	Mineral Concentrations Hardness Code
Horizon Transition Distinctness Code	Secondary Carbonates Code	Voids Size mm	Mineral Concentrations Hardness Desc
Horizon Transition Distinctness Desc	Secondary Carbonates Desc	Voids Abundance Class	Roots Abundance Code
Horizon Transition Distinctness cm	Colour Moist Munsell SCMO	Voids Abundance Desc	Roots Abundance Desc
Horizon Transition Topography Code	Hue Moist	Voids Abundance Number	Roots Abundance No
Horizon Transition Topography Desc	Value Moist	Coatings Nature Code	Roots Diameter Code
Structure Type Code	Chroma Moist	Coatings Nature Desc	Roots Diameter Desc
Structure Type Desc	Colour Moist Desc	Coatings Form Code	Roots Diameter mm
Structure Strength Code	Colour Dry Munsell SCDR	Coatings Form Desc	Biological Features Code
Structure Strength Desc	Hue Dry	Coatings Location Code	Biological Features Desc
Structure Size Code	Value Dry	Coatings Location Desc	Notes
Structure Size Desc	Chroma Dry	Cementation_Compaction Nature Code	
Coarse Fragments Abundance Code	Colour Dry Desc	Cementation_Compaction Nature Desc	

Horizons Lab Data

PRID	EC dS_m 2_5_susp EL25	Perc Exch Na	Cu
HONU	EC uS_cm Sat_Paste ELCO	CEC Effective cmol_kg from Sum of Exch Bases	S
Latitude	EC dS_m Sat_Paste ELCO	CEC Measured cmol_kg	Zn
Longitude	Ca perc Exch plus Sol Mehlich_3 ICP-AES	BS perc Estimate from pHwater	Fe_ox (oxalate extract)
Soil Lab ID	Mg perc Exch plus Sol Mehlich_3 ICP-AES	BS perc from CECeffective_div_CECmeasured	Al_ox (oxalate extract)
Horizon Code Short	K perc Exch plus Sol Mehlich_3 ICP-AES	Gypsum perc	inorgC
Horizon Code Original	Na perc Exch plus Sol Mehlich_3 ICP-AES	Gypsum g_kg GYPS	totC
Horizon Thickness cm	Ca mg_kg Extr 1M NH4acetate_AAS SOCA	Tot N mg_kg TOTN	orgC
Upper Depth cm	Mg mg_kg Extr 1M NH4acetate_AAS SOMG	Tot N g_kg TOTN Kjeldahl	exchH
Lower Depth cm	K mg_kg Extr 1M NH4acetate_AAS SOLK	Tot N perc	exchAl

Dominant Sand Grade Code	Na mg_kg Extr 1M NH4acetate_AAS SONA	Tot P mg_kg Exch plus Sol Mehlich_3 ICP-AES	exchAcid
Dominant Sand Grade Desc	Cl mg_kg Extr SOCL	Tot P mg_kg Mehlich_3 ICP-AES	totMo
Sand perc 53um SDTO	Chloride perc Extr	P Olsen mg_kg	totCd
Silt perc 2_53um STPC	Sol Sulfate ppm SSO4	P2O5 mg_kg Olsen Px2_291	totPb
Clay per 2um CLPC	Sulfate perc	Organic Carbon perc	totV
Texture Code PSCL	Sol Carbonate meq_l SCO3	Organic Carbon g_kg	totHg
Texture Desc	Tot Carbonate Equivalent g_kg TCEQ	Organic Matter perc OCx1_74	totCr
Silt plus Clay perc	Sol Bicarbonate meq_l HCO3	Mn mg_kg Exch plus Sol Mehlich_3 ICP-AES	totCo
Bulk Density kg_dm3	Exch Ca cmol_kg EXCA	Fe mg_kg Exch plus Sol Mehlich_3 ICP-AES	totNi
pH CaCL2 2_5_susp	Exch Mg cmol_kg EXMG	Al mg_kg Exch plus Sol Mehlich_3 ICP-AES	totCu
pH water	Exch K cmol_kg EXCK	Porosity	totZn
pH KCl	Exch Na cmol_kg EXNA	CaCO3	totAs
EC uS_cm 2_5_susp EL25	Perc Exch Na plus Exch Mg	B	totSb

G. NSPD MS Excel tables and fields.

Profile / Point Registration

PRID	Latitude	WRB2015 Reclassification RSG
PRID Let	Longitude	WRB2015 Reclassification Principal Qualifiers
PRID Dig	Elevation m GPS	WRB2015 Reclassification Supplementary Qualifiers
ALT PRID	Elevation m SRTM	WRB Classification by S Nambambi
Date (YYYY-MM-DD)	Country ID	Degree Square
Date (DD-Mon-YY)	Data Origin	Quarter Degree Square
Date (DD/MM/YYYY)	Surveyor(s)	Farm No
Year	Status	Farm Name
Month	Profile Classification Original	Location
Day	Profile Classification Original Code	Notes
Coordinate Reference System	Classification System	Photos

Profile / Point - Environmental Covariates Derived from Maps

PRID	NSS1M SOIL3	OKA Soil1
Latitude	NSS1M SOIL4	OKA Soil2
Longitude	NSS1M SOIL5	OKA Slope Description
AEZ 1998 Code	NSS1M SOIL6	OKA Texture
AEZ 1998 Description	NSS1M SOIL7	OKA Salinity
GPZ 1998 Code	NSS1M Dissection	OKA Sodicity
GPZ 1998 Description	NSS1M Relative Relief	Landscapes Atlas 2002
Catchment Strohbach	OWA Epigraph	Landscapes Short Atlas 2002
Final Catchment Strohbach	OWA LandUnit1	Land Surface Forms USGS-MORAP
Vegetation Type Giess	OWA LandUnit2	Landforms Iwahashi-Pike
CC 2003 kgAB/ha/a	OWA LandUnit3	Landforms Meybeck
CC 2021 kgAB/ha/a	OWA Soil Types	GEO Lithology Code
FHFA Code	OWA Soil1	GEO Main Lithology
FHFA Description	OWA Soil2	GEO Age
FHFA Potential	OWA Soil3	GEO Other Rocks
Slope% SRTM	OWA Soil4	GEO Supergroup
Ruggedness SRTM	OWA Slope Description	GEO Group
Aspect Deg SRTM	OWA Texture	GEO Subgroup
Soil Unit 2021 Coetzee	OWA Surface Fragments	GEO Formation
NSS1M Epigraph	OWA Salinity	GEO Complex
NSS Land Division	OWA Sodicity	GEO Suite
NSS Land Province	OKA Epigraph	GEO Intrusive
NSS Land Region	OKA LandUnit1	GEO Unit

NSS1M Soil Unit	OKA LandUnit2	GEO Lithology Class
NSS1M SOIL1	OKA LandUnit2	GEO Metamorphism
NSS1M SOIL2	OKA Soil Types	

Profile / Point Site Data

PRID	Human Influence Code	Surface Sealing Hardness Desc
Latitude	Human Influence Desc	Surface Cracks Width cm
Longitude	Rock Outcrops Abundance/Cover Code	Surface Cracks Width Code
Diagnostic Horizon / Property / Material	Rock Outcrops Cover%	Distance Between Surface Cracks (m)
Present Weather Code	Rock Outcrops Cover Desc	Salt (Surface Salinity)
Present Weather Desc	Rock Outcrops - m between	Bleached Sand Code
Past Weather Code	Rock Outcrops Type	Bleached Sand%
Past Weather Desc	Coarse Surface Fragments Cover Code	Bleached Sand Desc
Soil Temperature Regime Code	Coarse Surface Fragments Cover%	Flooded >2 weeks/a
Soil Temperature Regime Desc	Coarse Surface Fragments Cover Desc	Major Landform Code (2nd level)
Soil Moisture Regime Code	Coarse Surface Fragments Size Code	Major Landform Desc
Soil Moisture Regime Desc	Coarse Surface Fragments Size cm	Minor Landform desc
Slope Gradient Class	Coarse Surface Fragments Size Desc	Position code
Slope Form Vert	Surface Coarse Surface Fragments Type	Position Desc
Slope Form Hor	Erosion Type Code	Vegetation
Slope Gradient Measured%	Erosion Type	Land Cover
Slope Gradient Desc	Erosion Area Affected Code	Parent Material
Gradient% (for landform) from SRTM	Erosion Area Affected%	Lithology
Slope% SRTM	Erosion Degree Code	Drainage Desc
Relief m/km	Erosion Degree Desc	Drainage Code
Drainage density	Erosion Status Code	Geology
Present Land Use Code	Erosion Status Desc	Effective Depth
Present Land Use Desc	Surface Sealing Thickness Code	Depth Class
Past Land Use Code	Surface Sealing Thickness mm	Biological activity
Past Land Use Desc	Surface Sealing Thickness Desc	Notes
Grazer Type	Surface Sealing Hardness Code	

Horizon / Layer Field Data

PRID	Drainage Code	Voids Size mm
HONU	Drainage Desc	Voids Abundance Class
Latitude	Field Texture Code	Voids Abundance Desc
Longitude	Field Texture Desc	Voids Abundance Number
Representative Profile	Carbonate Estimate%	Coatings Nature Code
Diagnostic Horizon	Carbonate Reaction Code	Coatings Nature Desc
Diagnostic Property	Carbonate Reaction Desc	Coatings Form Code
Diagnostic Material	Secondary Carbonates Code	Coatings Form Desc
Qualifier	Secondary Carbonates Desc	Coatings Location Code
Horizon Code Short	Colour Moist [Munsell] (SCMO)	Coatings Location Desc
Horizon Code Original	Hue [Moist]	Cementation / Compaction Nature Code
Subordinate Horizon Code Original	Value [Moist]	Cementation / Compaction Nature Desc
Lithology of R/C	Chroma [Moist]	Cementation / Compaction Continuity Code
Horizon Thickness cm	Colour Moist Desc [Munsell]	Cementation / Compaction Continuity Desc
Upper Depth cm	Colour Dry [Munsell] (SCDR)	Cementation / Compaction Fabric Code
Lower Depth cm	Hue [Dry]	Cementation / Compaction Fabric Desc
Horizon Transition Distinctness Code	Value [Dry]	Mineral Concentrations Nature Code
Horizon Transition Distinctness Desc	Chroma [Dry]	Mineral Concentrations Nature Desc
Horizon Transition Distinctness cm	Colour Dry Desc [Munsell]	Mineral Concentrations Abundance Code
Horizon Transition Topography Code	Mottling Desc	Mineral Concentrations Abundance Desc
Horizon Transition Topography Desc	Vertic Properties	Mineral Concentrations Abundance%
Structure Type Code	Consistence Dry Code	Mineral Concentrations Size Code
Structure Type Desc	Consistence Dry Desc	Mineral Concentrations Size Desc
Structure Strength Code	Consistence Moist Code	Mineral Concentrations Shape Code
Structure Strength Desc	Consistence Moist Desc	Mineral Concentrations Shape Desc
Structure Size Code	Consistence Wet - Stickiness Code	Mineral Concentrations Hardness Code
Structure Size Desc	Consistence Wet - Stickiness Desc	Mineral Concentrations Hardness Desc
Coarse Fragments Abundance Code	Consistence Wet - Plasticity Code	Roots Abundance Code
Coarse Fragments Abundance Desc	Consistence Wet - Plasticity Desc	Roots Abundance Desc
Coarse Fragments Abundance%	Porosity Measured	Roots Abundance No
Coarse Fragments Size Code	Porosity Class Code	Roots Diameter Code
Coarse Fragments Size Desc	Porosity Class Desc	Roots Diameter Desc
Coarse Fragments Size mm	Porosity Class%	Roots Diameter mm
Coarse Fragments Shape Code	Voids Type Code	Biological Features Code
Coarse Fragments Shape Desc	Voids Type Desc	Biological Features Desc
Coarse Fragments Degree of Weathering	Voids Size Class	Notes
Coarse Fragments Type	Voids Size Desc	

Horizon / Layer Lab Data

PRID	Silt + Clay%
HONU	Bulk Density kg/dm ³
Latitude	pH CaCl ₂ (PHCA) [2 soil:5 CaCl ₂ suspension]
Longitude	pH water (PHAQ)
Laboratory	pH KCl (PHKC)
Soil Lab ID	Electrical Conductivity µS/cm [2 soil:5 water suspension] (EL25)
Horizon Code Short	Electrical Conductivity dS/m [2 soil:5 water suspension] (EL25)
Horizon Code Original	Electrical Conductivity µS/cm [Saturated Paste Extract] (ELCO)
Horizon Thickness cm	Electrical Conductivity dS/m [Saturated Paste Extract] (ELCO)
Upper Depth cm	Ca% [Exchangeable & Soluble; Mehlich #3, ICP-AES]
Lower Depth cm	Mg% [Exchangeable & Soluble; Mehlich #3, ICP-AES]
Dominant Sand Grade Code	K% [Exchangeable & Soluble; Mehlich #3, ICP-AES]
Dominant Sand Grade Desc	Na% [Exchangeable & Soluble; Mehlich #3, ICP-AES]
Sand% < 53 µm (SDTO)	Ca mg/kg [Extractable; 1M NH ₄ acetate, AAS] (SOCA)
Silt% [2-53µm] (STPC)	Mg mg/kg [Extractable; 1M NH ₄ acetate, AAS] (SOMG)
Clay% [< 2 µm] (CLPC)	K mg/kg [Extractable; 1M NH ₄ acetate, AAS] (SOLK)
Texture Code (PSCl)	Na mg/kg [Extractable; 1M NH ₄ acetate, AAS] (SONA)
Texture Desc	Cl mg/kg [Extractable] (SOCL)
Chloride% [Extractable]	Base Saturation% [CEC _{effective} /CEC _{measured}]
SSO ₄ ppm	Gypsum%
Sulfate%	Gypsum g/kg (GYPS)
SCO ₃ Soluble Carbonate meq/l	Total Nitrogen mg/kg (TOTN)
Total Carbonate Equivalent g/kg (TCEQ)	Total Nitrogen mg/kg (TOTN)
Soluble Bicarbonate meq/l (HCO ₃)	Total Nitrogen g/kg (TOTN)
Secondary Carbonates Code	Total Nitrogen%
Secondary Carbonates Desc	Total Phosphorus mg/kg [Mehlich #3, ICP-AES]
Exchangeable Calcium cmol(+)/kg (EXCA)	Total Phosphorus unit? [Mehlich #3, ICP-AES]
Exchangeable Magnesium cmol(+)/kg (EXMG)	Phosphorus [Olsen] mg/kg
Exchangeable Potassium cmol(+)/kg (EXCK)	P ₂ O ₅ mg/kg [Olsen; P x 2.291]
Exchangeable Sodium cmol(+)/kg (EXNA)	Organic Carbon%
% (ExNa+ExMg)	Organic Carbon g/kg
% ExNa	Organic Matter% [OC*1.74]
CEC Effective cmol(+)/kg [Sum of Exchangeable Bases]	Mn mg/kg [Exchangeable & Soluble; Mehlich #3, ICP-AES]
CEC Measured cmol(+)/kg	Fe mg/kg [Exchangeable & Soluble; Mehlich #3, ICP-AES]
Base Saturation Estimate% [from pH _{water}]	Al mg/kg [Exchangeable & Soluble; Mehlich #3, ICP-AES]

H. Soil typological units identified in Namibia.

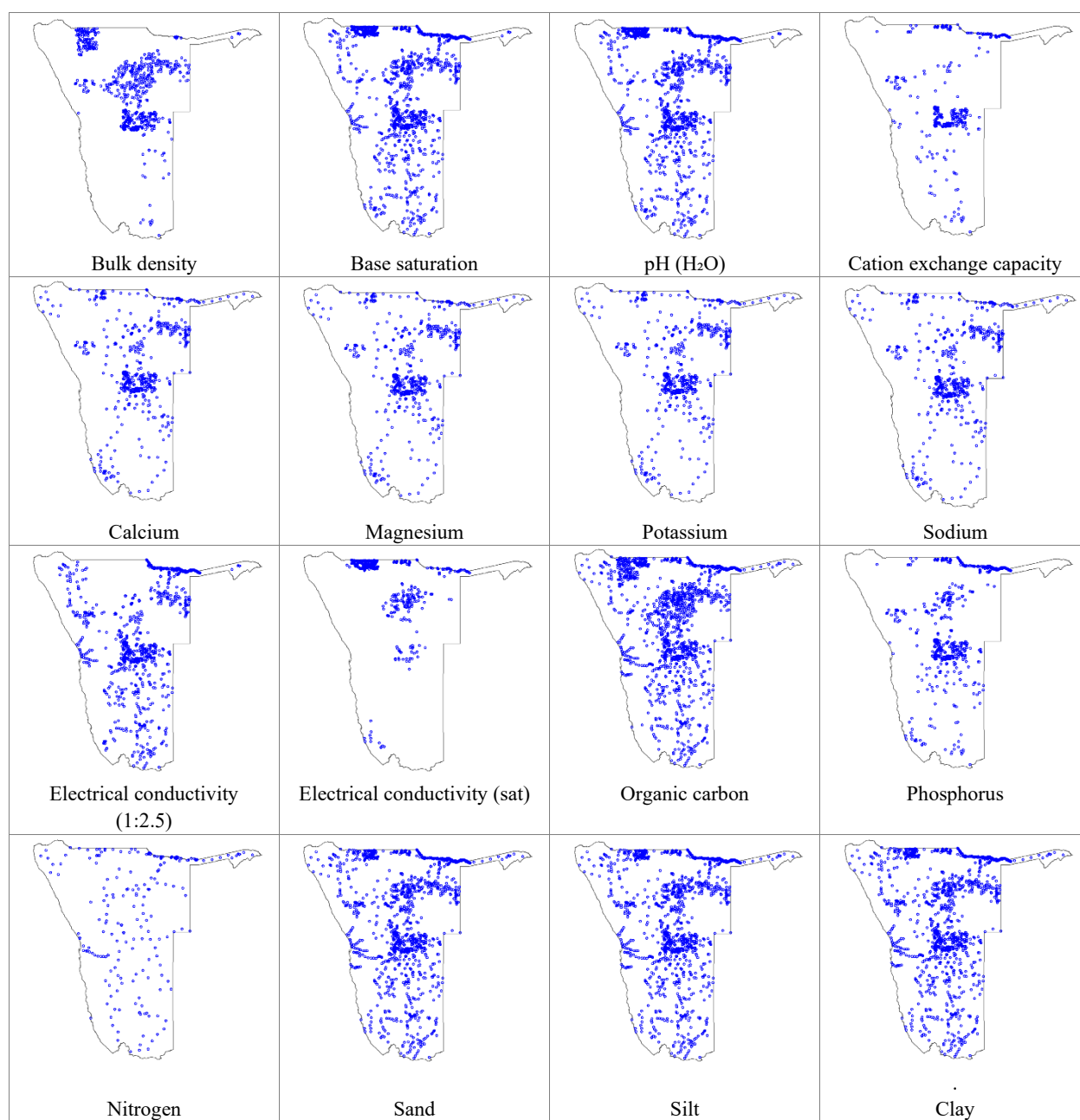
RSG	RSG and highest-ranking principal qualifier	Full classification
Arenosol	Arenosol	Arenosol
		Arenosol (Aeolic)
	Calcaric Arenosol	Calcaric Arenosol
	Chromic Arenosol	Chromic Arenosol
		Chromic Arenosol (Protocalcic)
		Calcaric Chromic Arenosol
	Endosalic Arenosol	Calcaric Sodic Endosalic Arenosol
		Fluvic Endosalic Arenosol
		Sodic Endosalic Arenosol
	Eutric Arenosol	Eutric Arenosol
	Fluvic Arenosol	Fluvic Arenosol
	Protic Arenosol	Protic Arenosol (Aeolic)
	Protosalic Arenosol	Fluvic Protosalic Arenosol
	Sideralic Arenosol	Sideralic Arenosol
		Sideralic Arenosol (Protocalcic)
		Calcaric Chromic Sideralic Arenosol
		Calcaric Sideralic Arenosol
		Chromic Sideralic Arenosol
		Fluvic Sideralic Arenosol
		Fluvic Sideralic Arenosol (Protocalcic)
		Protoargic Sideralic Arenosol
		Sodic Salic Sideralic Arenosol
Calcisol	Cambic Calcisol	Cambic Calcisol
		Cambic Calcisol (Arenic)
		Cambic Calcisol (Arenic, Endosalic, Sodic)
		Skeletal Cambic Calcisol
	Haplic Calcisol	Haplic Calcisol
		Haplic Calcisol (Arenic)
		Haplic Calcisol (Hypercalcic)
		Haplic Calcisol (Hypercalcic) (Yermic)
		Haplic Calcisol (Hypercalcic, Endosalic)
		Haplic Calcisol (Loamic)
		Haplic Calcisol (Loamic, Fluvic)
		Haplic Calcisol (Sodic)
		Haplic Calcisol (Yermic)
	Leptic Calcisol	Skeletal Cambic Leptic Calcisol
		Skeletal Leptic Calcisol
	Luvic Calcisol	Luvic Calcisol
		Cambic Luvic Calcisol (Sodic)
	Petric Calcisol	Petric Calcisol
		Petric Calcisol (Arenic)
		Petric Calcisol (Fluvic)
		Petric Calcisol (Fluvic, Yermic)
		Petric Calcisol (Hypercalcic)
		Petric Calcisol (Hypercalcic, Sodic)
		Petric Calcisol (Sodic)
		Petric Calcisol (Yermic)
		Gypsic Petric Calcisol
		Gypsic Petric Calcisol (Fluvic, Yermic)
		Gypsic Petric Calcisol (Yermic)
		Leptic Petric Calcisol
	Skeletal Calcisol	Skeletal Calcisol (Hypercalcic)
Cambisol	Cambisol	Cambisol (Arenic)
		Cambisol (Arenic, Yermic)
	Calcaric Cambisol	Calcaric Cambisol
	Chromic Cambisol	Chromic Cambisol
		Chromic Cambisol (Arenic)
		Calcaric Chromic Cambisol

		Dolomitic Chromic Cambisol
		Eutric Chromic Cambisol
		Leptic Chromic Cambisol
	Eutric Cambisol	Eutric Cambisol
		Eutric Cambisol (Protocalcic, Yermic)
	Fluvic Cambisol	Fluvic Cambisol
		Fluvic Cambisol (Arenic)
		Fluvic Cambisol (Loamic)
		Eutric Fluvic Cambisol (Arenic)
		Eutric Fluvic Cambisol (Loamic)
	Leptic Cambisol	Leptic Cambisol
		Calcaric Leptic Cambisol
		Chromic Leptic Cambisol
		Dolomitic Chromic Leptic Cambisol
		Dolomitic Leptic Cambisol
		Eutric Chromic Leptic Cambisol
		Eutric Leptic Cambisol
		Salic Vertic Leptic Cambisol
		Skeletal Leptic Cambisol
	Salic Cambisol	Calcaric Sodic Salic Cambisol
		Sodic Salic Cambisol
		Sodic Salic Cambisol (Arenic)
	Sideralic Cambisol	Sideralic Cambisol (Arenic)
		Calcaric Sideralic Cambisol
		Fluvic Sideralic Cambisol (Arenic)
	Skeletal Cambisol	Eutric Skeletal Cambisol
	Sodic Cambisol	Sodic Cambisol
		Sodic Cambisol (Arenic)
		Sodic Cambisol (Loamic)
	Vertic Cambisol	Vertic Cambisol
		Vertic Cambisol (Arenic)
Durisol	Calcic Durisol	Calcic Durisol
		Calcic Durisol (Arenic)
		Leptic Calcic Durisol
	Gypsic Durisol	Calcic Gypsic Durisol
	Petric Durisol	Petric Durisol
		Calcic Petric Durisol
		Leptic Calcic Petric Durisol
		isDA calcic Petric Durisol
	Petrocalcic Durisol	Petrocalcic Durisol
Fluvisol	Fluvisol	Fluvisol (Arenic, Protocalcic)
		Fluvisol (Clayic)
		Fluvisol (Loamic)
		Fluvisol (Loamic, Protocalcic)
	Calcaric Fluvisol	Calcaric Fluvisol
		Calcaric Fluvisol (Arenic)
		Calcaric Fluvisol (Humic)
		Calcaric Fluvisol (Yermic)
		Eutric Calcaric Fluvisol
		Skeletal Calcaric Fluvisol
	Eutric Fluvisol	Eutric Fluvisol
		Eutric Fluvisol (Arenic)
	Gleyic Fluvisol	Gleyic Fluvisol
		Gleyic Fluvisol (Arenic)
		Gleyic Fluvisol (Loamic)
		Calcaric Gleyic Fluvisol (Arenic)
		Sodic Gleyic Fluvisol
	Gypsic Fluvisol	Calcaric Gypsic Fluvisol
		Calcaric Gypsic Fluvisol (Yermic)
	Leptic Fluvisol	Calcaric Leptic Fluvisol
		Calcaric Leptic Fluvisol (Arenic)

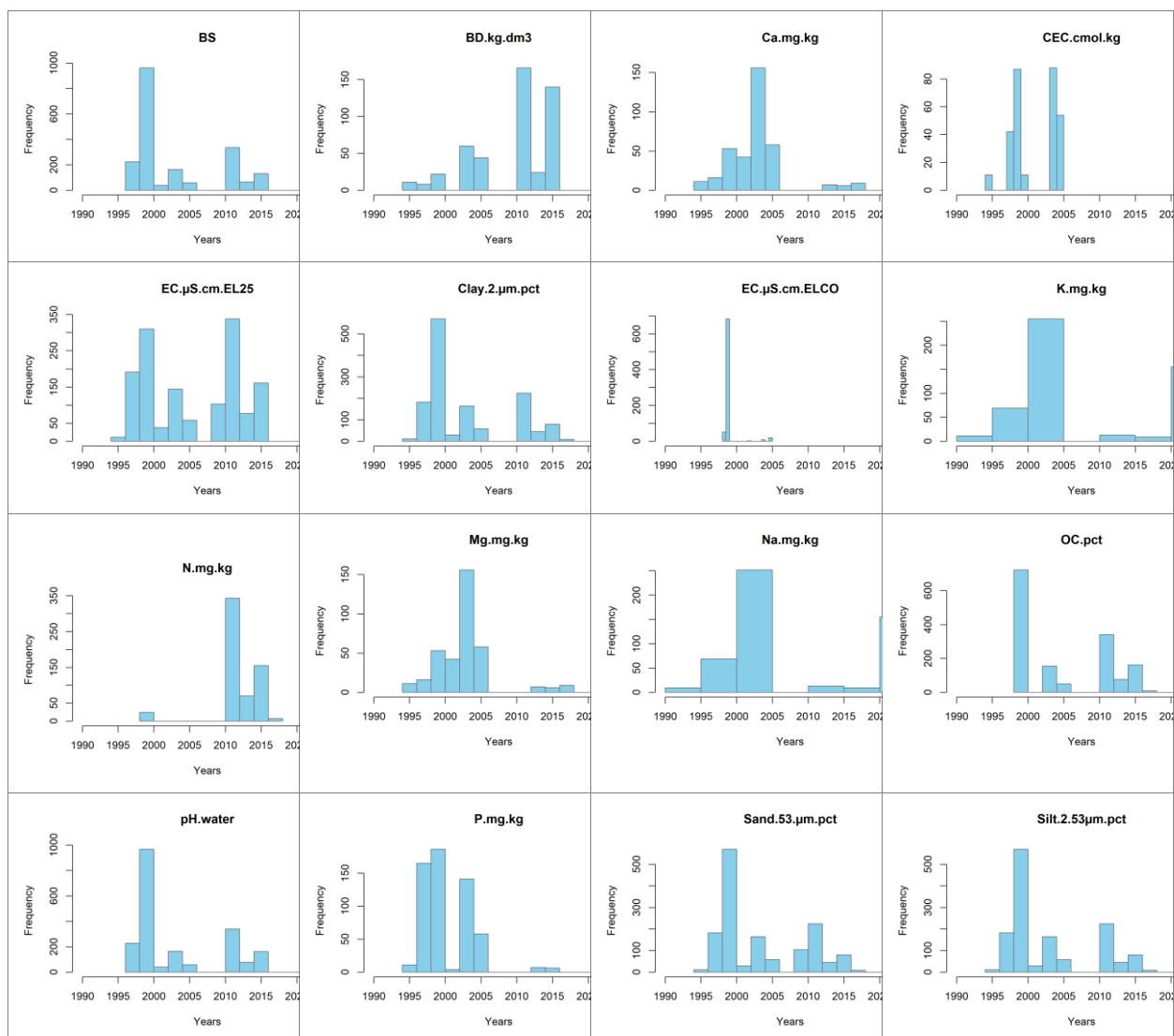
	Tidalic Fluvisol	Tidalic Fluvisol (Loamic, Sulfuric)
Gleysol	Fluvic Gleysol	Fluvic Gleysol
		Fluvic Gleysol (Arenic)
		Fluvic Gleysol (Clayic)
		Fluvic Gleysol (Siltic)
Gypsisol	Calcic Gypsisol	Calcic Gypsisol
		Calcic Gypsisol (Fluvic, Yermic)
		Calcic Gypsisol (Yermic)
	Haplic Gypsisol	Haplic Gypsisol
		Haplic Gypsisol (Yermic)
	Petric Gypsisol	Petric Gypsisol
		Petric Gypsisol (Yermic)
		Calcic Petric Gypsisol
		Calcic Petric Gypsisol (Fluvic)
		Calcic Petric Gypsisol (Fluvic, Yermic)
		Calcic Petric Gypsisol (Yermic)
Leptosol	Calcaric Leptosol	Calcaric Leptosol
		Eutric Calcaric Leptosol
		Eutric Calcaric Leptosol (Arenic)
	Cambic Leptosol	Cambic Leptosol
	Dolomitic Leptosol	Dolomitic Leptosol
	Eutric Leptosol	Eutric Leptosol
		Eutric Leptosol (Yermic)
	Hyperskeletal Leptosol	Tidalic Hyperskeletal Leptosol
	Lithic Leptosol	Lithic Leptosol
		Lithic Leptosol (Salic, Sodic)
		Lithic Leptosol (Yermic)
		Calcaric Lithic Leptosol
		Calcaric Lithic Leptosol (Arenic)
		Dolomitic Lithic Leptosol
		Eutric Skeletic Lithic Leptosol
		Hyperskeletal Lithic Leptosol
		Skeletic Lithic Leptosol
		Skeletic Lithic Leptosol (Yermic)
	Mollic Leptosol	Mollic Leptosol
	Nudilithic Leptosol	Nudilithic Leptosol
		Tidalic Nudilithic Leptosol
	Skeletic Leptosol	Skeletic Leptosol
		Dolomitic Skeletic Leptosol
		Eutric Skeletic Leptosol
Luvisol	Haplic Luvisol	Haplic Luvisol
Regosol	Calcaric Regosol	Calcaric Regosol
		Calcaric Regosol (Yermic)
	Dolomitic Regosol	Dolomitic Regosol
	Eutric Regosol	Eutric Regosol
		Eutric Regosol (Arenic)
	Leptic Regosol	Leptic Regosol
		Leptic Regosol (Arenic)
		Leptic Regosol (Arenic)
		Calcaric Gypsic Leptic Regosol (Yermic)
		Calcaric Leptic Regosol
		Calcaric Leptic Regosol (Arenic)
		Calcaric Skeletic Leptic Regosol
		Colluvic Skeletic Leptic Regosol
		Colluvic Skeletic Leptic Regosol (Yermic)
		Dolomitic Leptic Regosol
		Dolomitic Skeletic Leptic Regosol
		Dystric Leptic Regosol (Arenic)
		Eutric Leptic Regosol
		Eutric Skeletic Leptic Regosol
		Skeletic Leptic Regosol

	Skeletal Regosol	Skeletal Leptic Regosol (Arenic)
		Skeletal Leptic Regosol (Yermic)
		Skeletal Regosol
		Calcaric Skeletic Regosol
		Colluvic Skeletic Regosol
		Dolomitic Colluvic Skeletic Regosol
		Eutric Colluvic Skeletic Regosol
		Eutric Skeletic Regosol
		Eutric Skeletic Regosol (Yermic)
Solonchak	Fluvic Solonchak	Fluvic Solonchak
		Fluvic Solonchak (Arenic)
	Gleyic Solonchak	Fluvic Calcic Gleyic Solonchak (Arenic)
		Fluvic Calcic Sodic Gleyic Solonchak
	Petrosalic Solonchak	Sodic Gleyic Petrosalic Solonchak
	Sodic Solonchak	Calcic Sodic Solonchak
		Fluvic Sodic Solonchak
		Petrocalcic Sodic Solonchak
Solonetz	Calcic Solonetz	Calcic Solonetz
		Calcic Solonetz (Clayic)
	Gleyic Solonetz	Calcic Salic Gleyic Solonetz
	Salic Solonetz	Calcic Salic Solonetz
	Stagnic Solonetz	Calcic Stagnic Solonetz
		Petrocalcic Salic Stagnic Solonetz
		Salic Stagnic Solonetz
Technosol	Technosol	Technosol
Vertisol	Vertisol	Vertisol (Arenic, Calcaric)
		Vertisol (Calcaric)
	Haplic Vertisol	Haplic Vertisol
	Petrocalcic Vertisol	Petrocalcic Vertisol (Mollic)
	Sodic Vertisol	Calcic Gypsic Sodic Vertisol

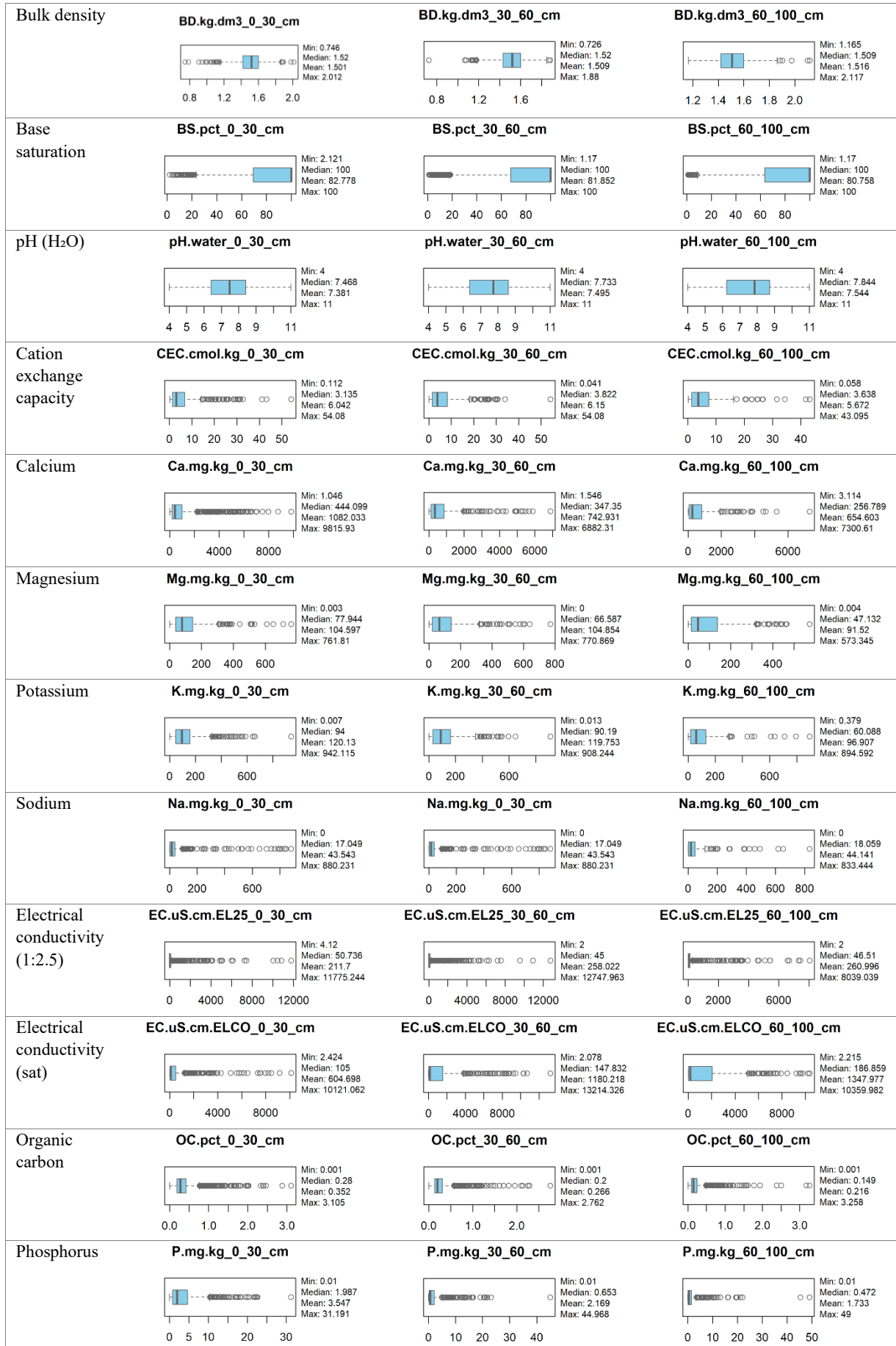
I. Spatial distribution of soil observations used for modelling, per soil property

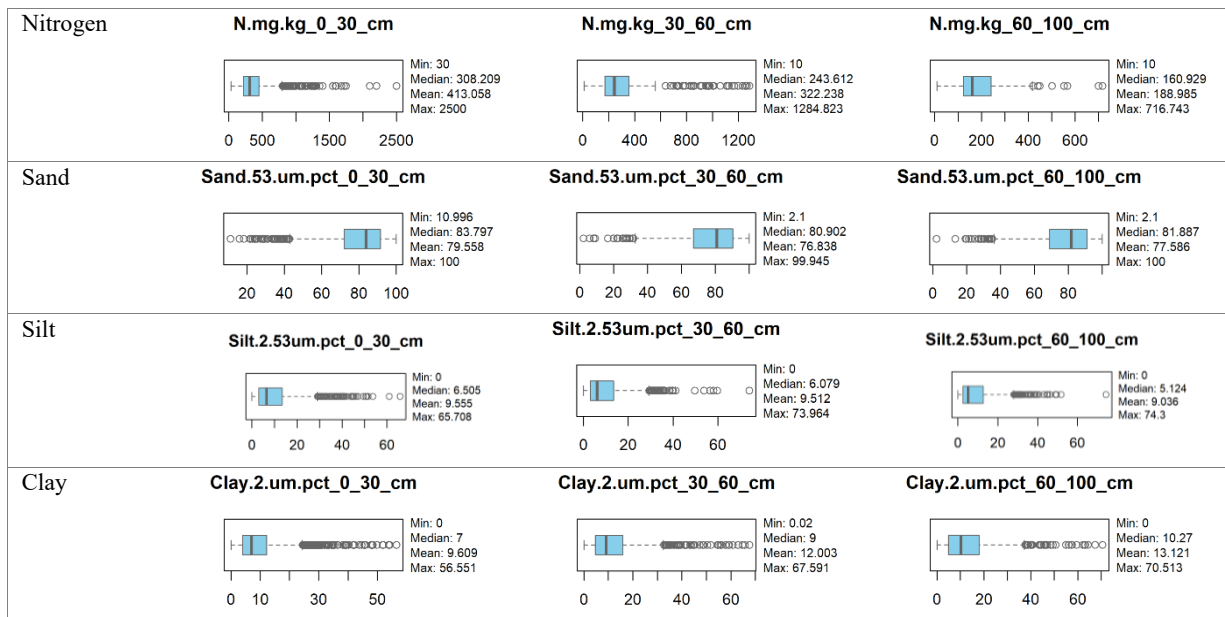


J. Temporal distribution of soil data

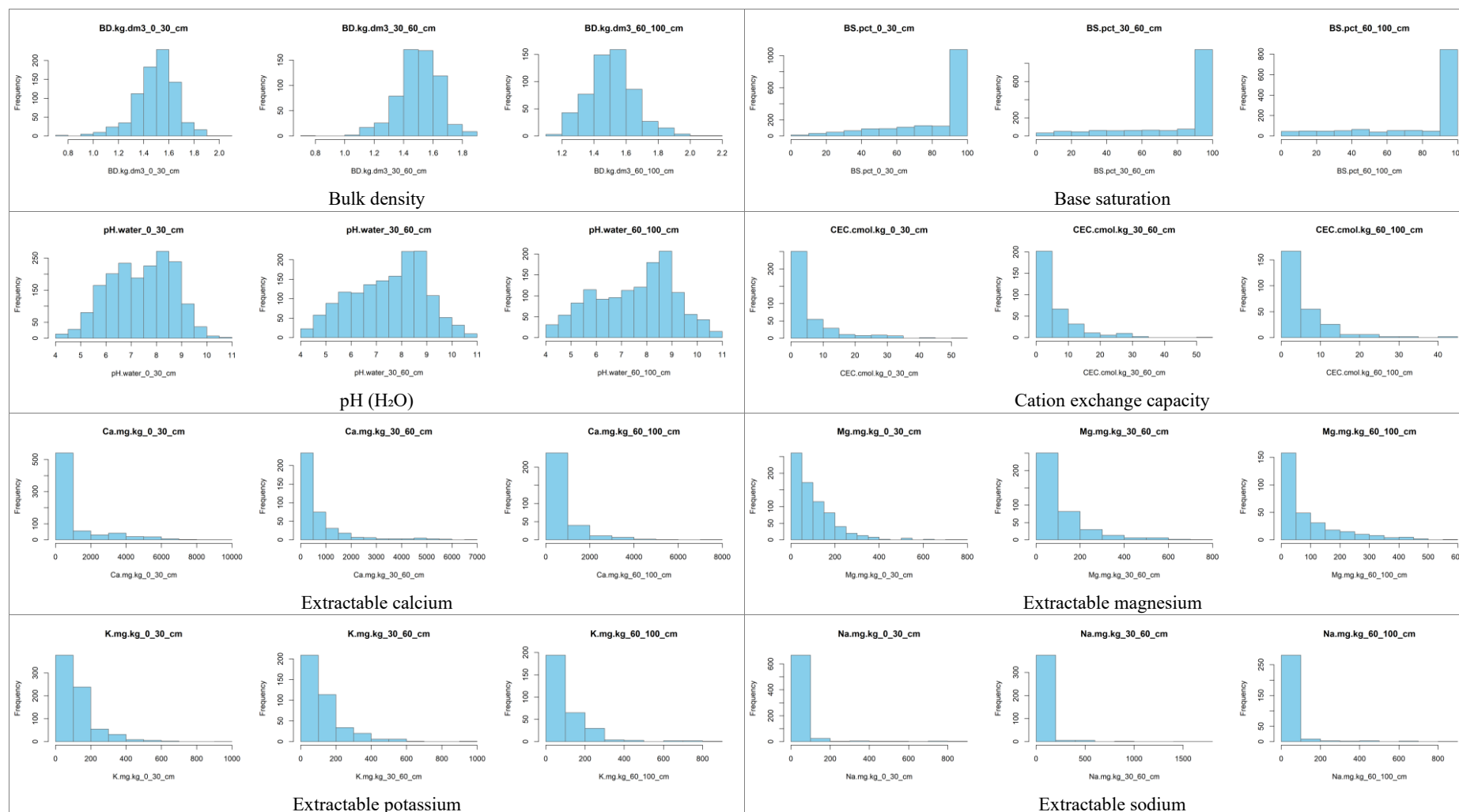


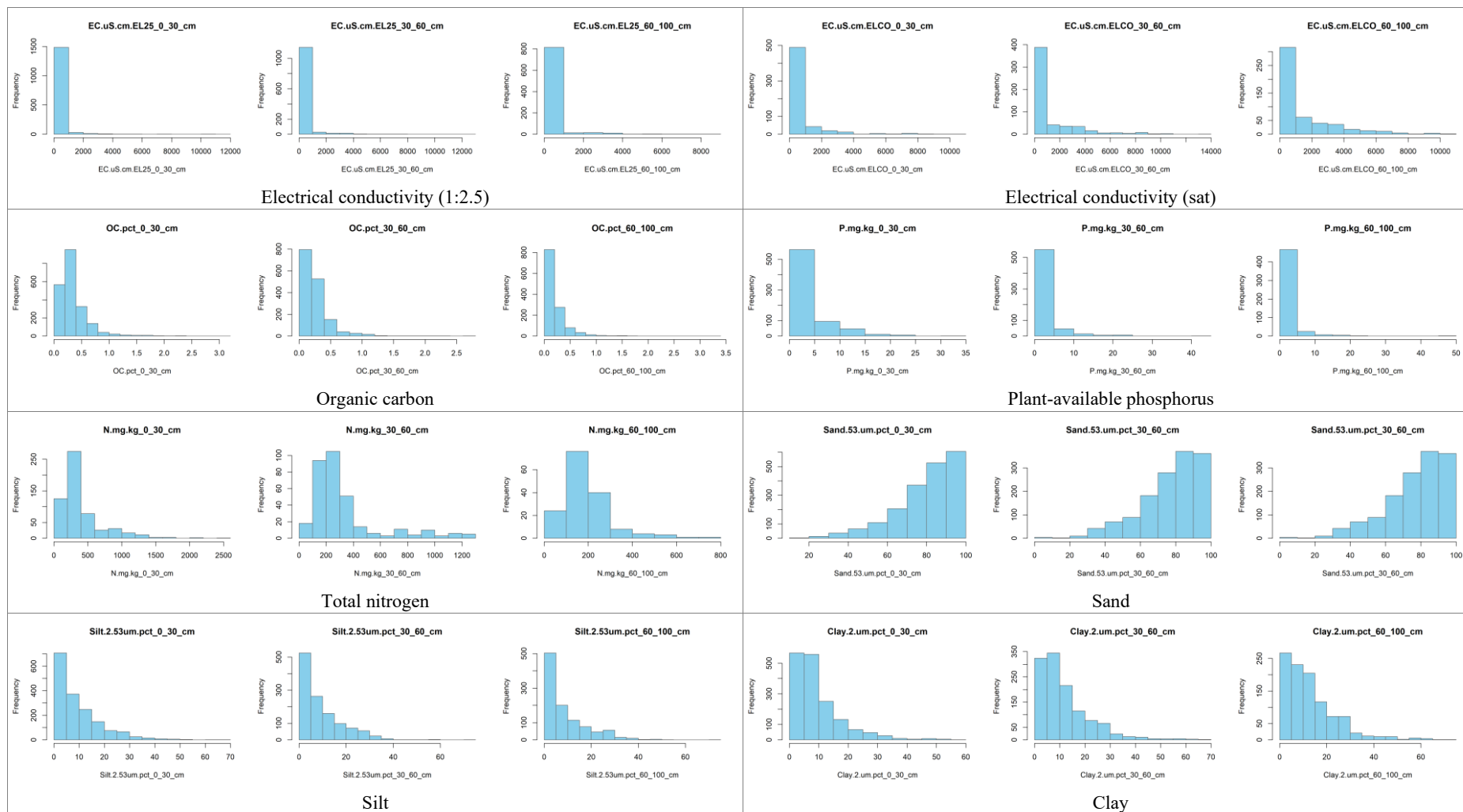
K. Boxplots of depth-harmonised soil property data





L. Histograms of depth-harmonised legacy soil property data.





Electrical conductivity (1:2.5)

Electrical conductivity (sat)

Organic carbon

Plant-available phosphorus

Total nitrogen

Sand

Silt

Clay

M. Environmental variables included in RF models — all properties and depths

[illegible]

N. Random Forests hyperparameters optimised for each soil property and depth interval.

mtry	BD			BS			pH (H ₂ O)			Cation Exchange Capacity		
	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm
	16	16	9	22	14	12	34	13	12	5	5	7
	7	6	2	4	12	7	8	11	2	3	3	2
nodesize												
sampsize	0.48	0.43	0.62	0.63	0.56	0.62	0.50	0.57	0.53	0.48	0.48	0.43

mtry	Ca			Mg			K			Sodium		
	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm
	16	12	5	10	4	1	8	1	1	1	6	6
	5	13	4	6	2	3	2	2	2	25	10	18
nodesize												
sampsize	0.52	0.73	0.60	0.56	0.55	0.73	0.65	0.82	0.82	0.26	0.54	0.88

mtry	EC (1:2.5)			EC (sat)			OC			P		
	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm
	1	1	1	16	16	9	10	18	1	4	4	4
	2	2	2	9	9	15	4	4	4	4	4	4
nodesize												
sampsize	0.33	0.33	0.33	0.56	0.56	0.78	0.65	0.42	0.50	0.66	0.66	0.66

mtry	N			Sand			Silt			Clay		
	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm	0–30 cm	30–60 cm	60–100 cm
	1	17	2	18	41	28	25	11	23	14	23	17
	2	6	6	2	11	8	7	4	9	4	15	14
nodesize												
sampsize	0.62	0.41	0.48	0.47	0.52	0.53	0.68	0.68	0.52	0.50	0.55	0.55

(*ntree* = 150 and *max_Nodes* = null for all properties and depths)

O. Descriptive statistics of performance metrics across all bootstrap iterations

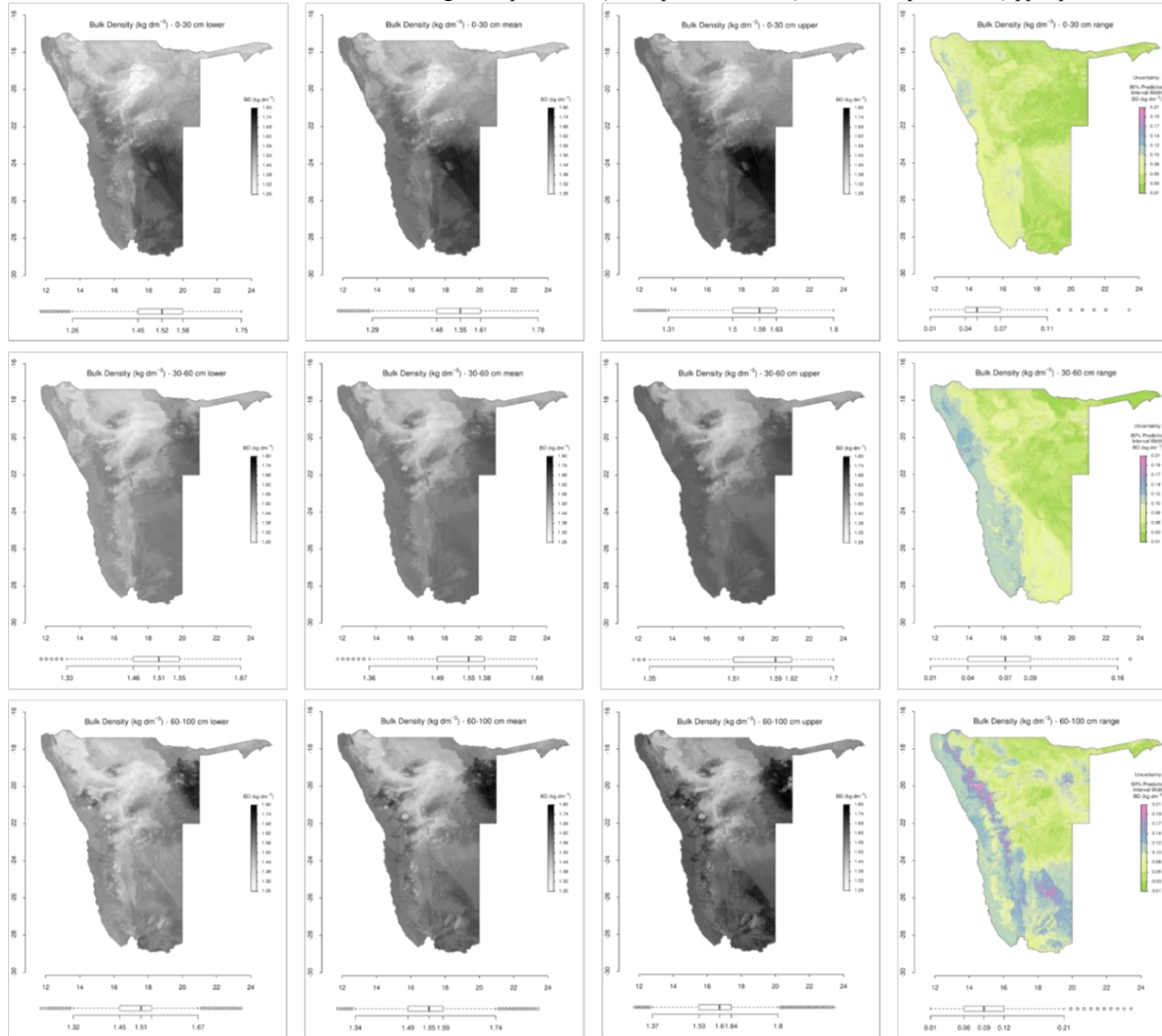
	Bulk density (kg dm ⁻³)						Base saturation (%)					
	0–30 cm		30–60 cm		60–100 cm		0–30 cm		30–60 cm		60–100 cm	
	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE
Median	0.38	0.13	0.19	0.12	0.31	0.12	0.44	18.2	0.45	21.3	0.46	21.5
Mean	0.37	0.13	0.21	0.12	0.29	0.13	0.45	18.3	0.44	21.3	0.46	21.7
Min	0.26	0.11	0.13	0.10	0.09	0.11	0.37	16.2	0.34	19.7	0.34	20.2
Max	0.52	0.14	0.29	0.14	0.39	0.14	0.55	20.4	0.52	24.0	0.55	23.4
SD	0.07	0.01	0.05	0.01	0.08	0.01	0.05	1.0	0.05	1.2	0.04	1.0
CV	19.3	6.6	22.1	8.9	27.8	6.8	12.0	5.3	12.4	5.7	9.3	4.7
	pH (H ₂ O)						Cation exchange capacity (cmol(+) kg ⁻¹)					
	0–30 cm		30–60 cm		60–100 cm		0–30 cm		30–60 cm		60–100 cm	
	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE
Median	0.51	0.87	0.46	1.08	0.48	1.13	0.46	5.5	0.49	5.1	0.16	5.5
Mean	0.52	0.87	0.46	1.07	0.49	1.13	0.46	5.5	0.49	5.1	0.19	6.0
Min	0.45	0.79	0.37	0.99	0.40	1.03	0.25	4.2	0.20	3.4	0.08	4.2
Max	0.57	0.95	0.54	1.13	0.55	1.23	0.68	6.8	0.68	6.5	0.41	8.4
SD	0.03	0.04	0.04	0.04	0.04	0.05	0.11	0.64	0.12	0.76	0.09	1.23
CV	5.7	4.3	9.3	3.9	9.1	4.5	25.1	11.8	24.0	14.9	47.5	20.6
	Calcium (mg kg ⁻¹)						Magnesium (mg kg ⁻¹)					
	0–30 cm		30–60 cm		60–100 cm		0–30 cm		30–60 cm		60–100 cm	
	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE
Median	0.45	1077	0.42	832	0.31	810	0.32	76.8	0.27	102.1	0.29	91.8
Mean	0.46	1081	0.45	830	0.32	836	0.31	82.2	0.27	101.8	0.30	90.3
Min	0.28	899	0.30	596	0.17	570	0.15	61.8	0.10	75.1	0.14	66.5
Max	0.59	1204	0.70	1047	0.48	1113	0.42	108.5	0.43	123.7	0.62	106.3
SD	0.1	86.7	0.1	126.3	0.1	141.6	0.08	13.7	0.08	12.9	0.12	11.0
CV	15.9	8.0	26.4	15.2	27.3	17.0	25.6	16.7	31.2	12.7	38.2	12.1
	Potassium (mg kg ⁻¹)						Sodium (mg kg ⁻¹)					
	0–30 cm		30–60 cm		60–100 cm		0–30 cm		30–60 cm		60–100 cm	
	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE
Median	0.41	84.0	0.41	89.5	0.36	95.0	0.10	95.3	0.14	110.4	0.18	78.0
Mean	0.40	87.0	0.42	91.9	0.32	96.0	0.11	95.7	0.15	114.2	0.22	76.9
Min	0.28	68.5	0.26	66.0	0.12	59.4	0.04	51.3	0.04	44.7	0.04	30.0
Max	0.52	131.4	0.61	141.4	0.48	171.8	0.22	140.1	0.61	203.1	0.50	152.2
SD	0.07	15.0	0.09	16.7	0.11	29.3	0.05	26.0	0.13	49.6	0.12	29.8
CV	17.1	17.2	20.3	18.2	33.6	30.5	43.9	27.2	85.2	43.4	55.0	38.8
	Electrical conductivity (1:2.5) (μS cm ⁻¹)						Electrical conductivity (sat) (μS cm ⁻¹)					
	0–30 cm		30–60 cm		60–100 cm		0–30 cm		30–60 cm		60–100 cm	
	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE
Median	0.37	541	0.20	766	0.10	750	0.14	1270	0.37	1572	0.41	1519
Mean	0.35	614	0.19	792	0.10	754	0.13	1266	0.37	1602	0.39	1512
Min	0.13	325	0.05	542	0.02	548	0.04	813	0.24	1300	0.20	1048
Max	0.59	919	0.35	1150	0.19	916	0.20	1589	0.50	2213	0.53	1789
SD	0.15	196.0	0.08	155.1	0.04	109.0	0.04	197.0	0.07	240.6	0.10	174.2
CV	41.5	31.9	44.6	19.6	41.	14.5	31.5	15.6	19.6	15.0	25.4	11.5

	Organic carbon (%)						Phosphorus (mg kg ⁻¹)					
	0–30 cm		30–60 cm		60–100 cm		0–30 cm		30–60 cm		60–100 cm	
	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE
Median	0.32	0.24	0.18	0.21	0.16	0.22	0.32	3.59	0.29	3.38	0.19	3.19
Mean	0.31	0.24	0.18	0.21	0.15	0.22	0.31	3.59	0.27	3.50	0.19	3.74
Min	0.19	0.20	0.10	0.17	0.07	0.15	0.14	2.79	0.07	2.33	0.02	1.78
Max	0.41	0.28	0.25	0.25	0.28	0.30	0.41	4.43	0.44	4.82	0.36	6.92
SD	0.06	0.03	0.04	0.03	0.06	0.05	0.06	0.41	0.10	0.78	0.11	1.38
CV	18.2	11.2	22.2	13.1	37.1	22.7	19.3	11.4	37.5	22.3	58.7	36.8
	Nitrogen (mg kg ⁻¹)						Sand (%)					
	0–30 cm		30–60 cm		60–100 cm		0–30 cm		30–60 cm		60–100 cm	
	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE
Median	0.56	210.1	0.76	129.6	0.28	104.0	0.45	11.7	0.37	13.4	0.33	13.9
Mean	0.57	211.3	0.75	131.7	0.25	101.5	0.45	11.6	0.36	13.3	0.33	13.9
Min	0.41	146.6	0.62	90.0	0.06	74.5	0.39	10.6	0.27	11.7	0.24	12.5
Max	0.77	276.5	0.86	165.6	0.45	144.1	0.50	12.9	0.44	14.4	0.45	15.0
SD	0.09	35.3	0.07	17.2	0.11	17.7	0.04	0.6	0.04	0.7	0.06	0.8
CV	16.5	16.7	9.7	13.1	45.3	17.5	7.9	5.3	11.2	5.5	17.0	5.9
	Silt (%)						Clay (%)					
	0–30 cm		30–60 cm		60–100 cm		0–30 cm		30–60 cm		0–30 cm	
	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE	R ²	RMSE
Median	0.43	7.0	0.39	7.1	0.36	7.3	0.31	6.9	0.23	9.1	0.28	9.3
Mean	0.44	7.0	0.39	7.2	0.34	7.5	0.32	7.0	0.22	9.2	0.28	9.1
Min	0.32	6.3	0.25	5.9	0.25	6.6	0.25	6.1	0.15	8.2	0.19	7.6
Max	0.54	8.3	0.48	8.7	0.44	8.5	0.42	8.12	0.30	10.5	0.42	10.4
SD	0.05	0.5	0.05	0.8	0.05	0.5	0.05	0.5	0.04	0.5	0.06	0.7
CV	11.3	7.1	13.9	10.7	15.3	6.8	15.1	7.5	17.4	5.7	20.	8.0

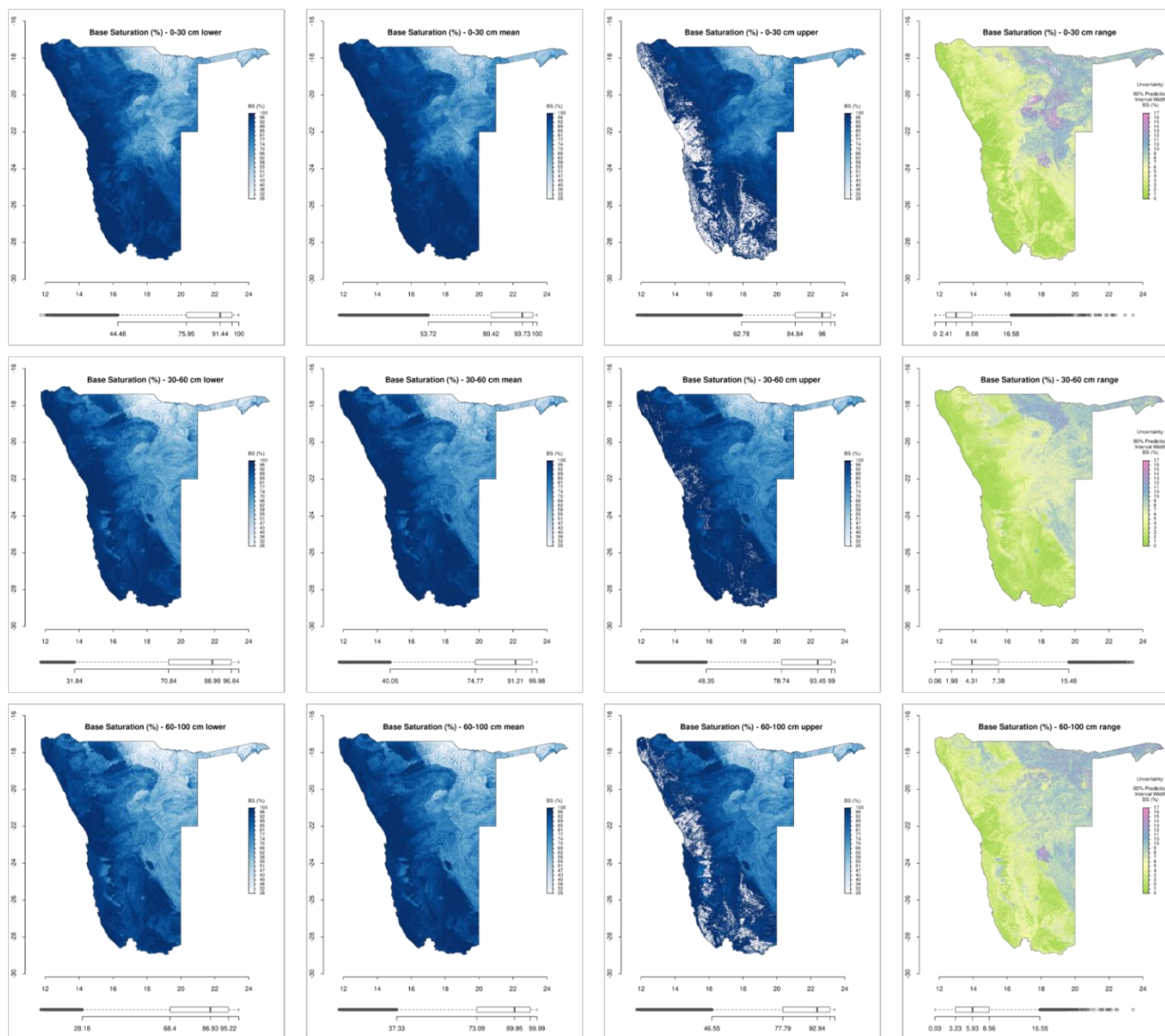
P. Spatial maps of prediction uncertainty for all soil properties and depth intervals

Top to bottom: depth classes 0–30 cm, 30–60 cm, 60–100 cm. Left to right: 5th percentile (lower prediction limit), mean, 95th percentile (upper prediction limit); 90% prediction interval

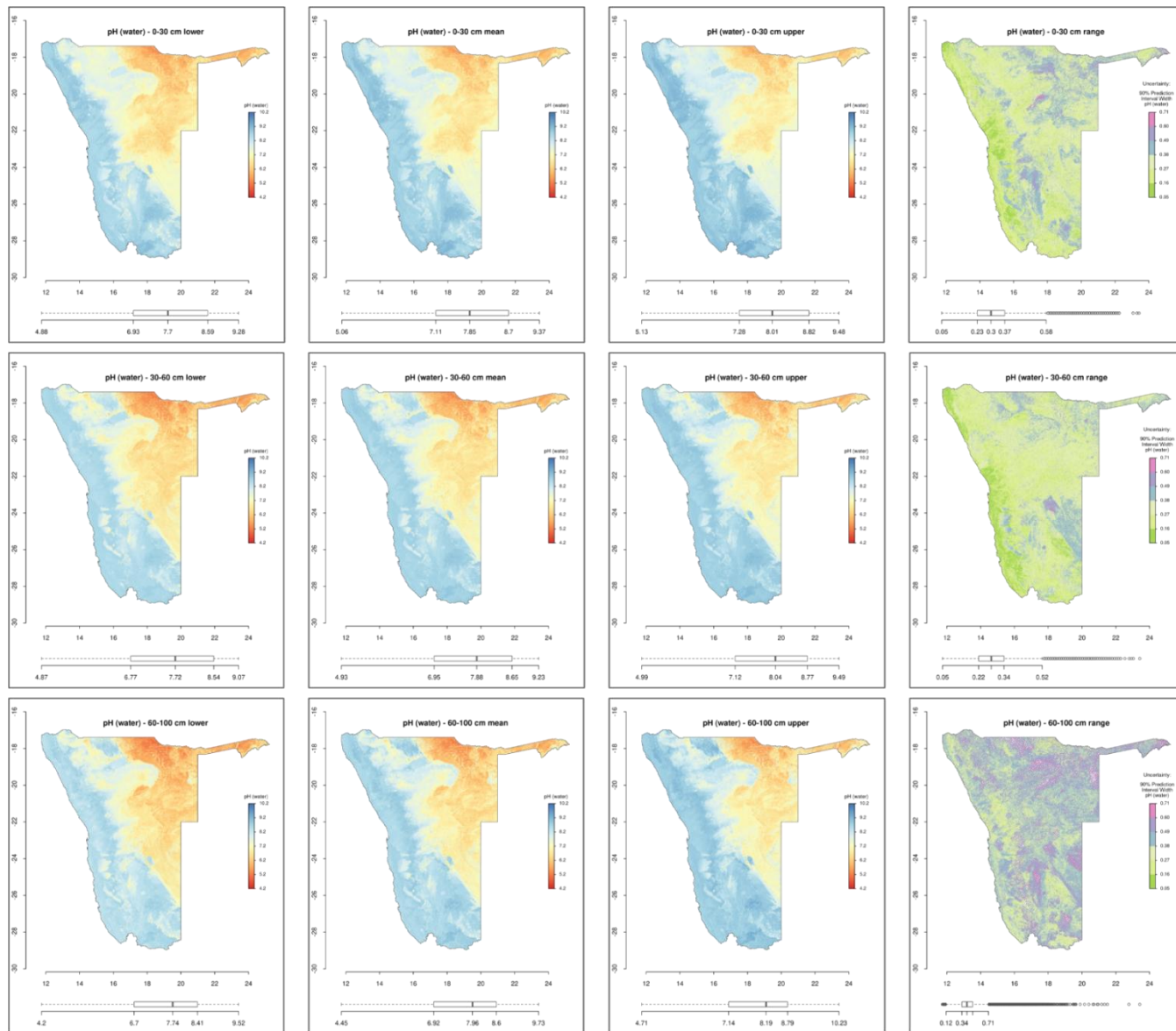
Bulk density



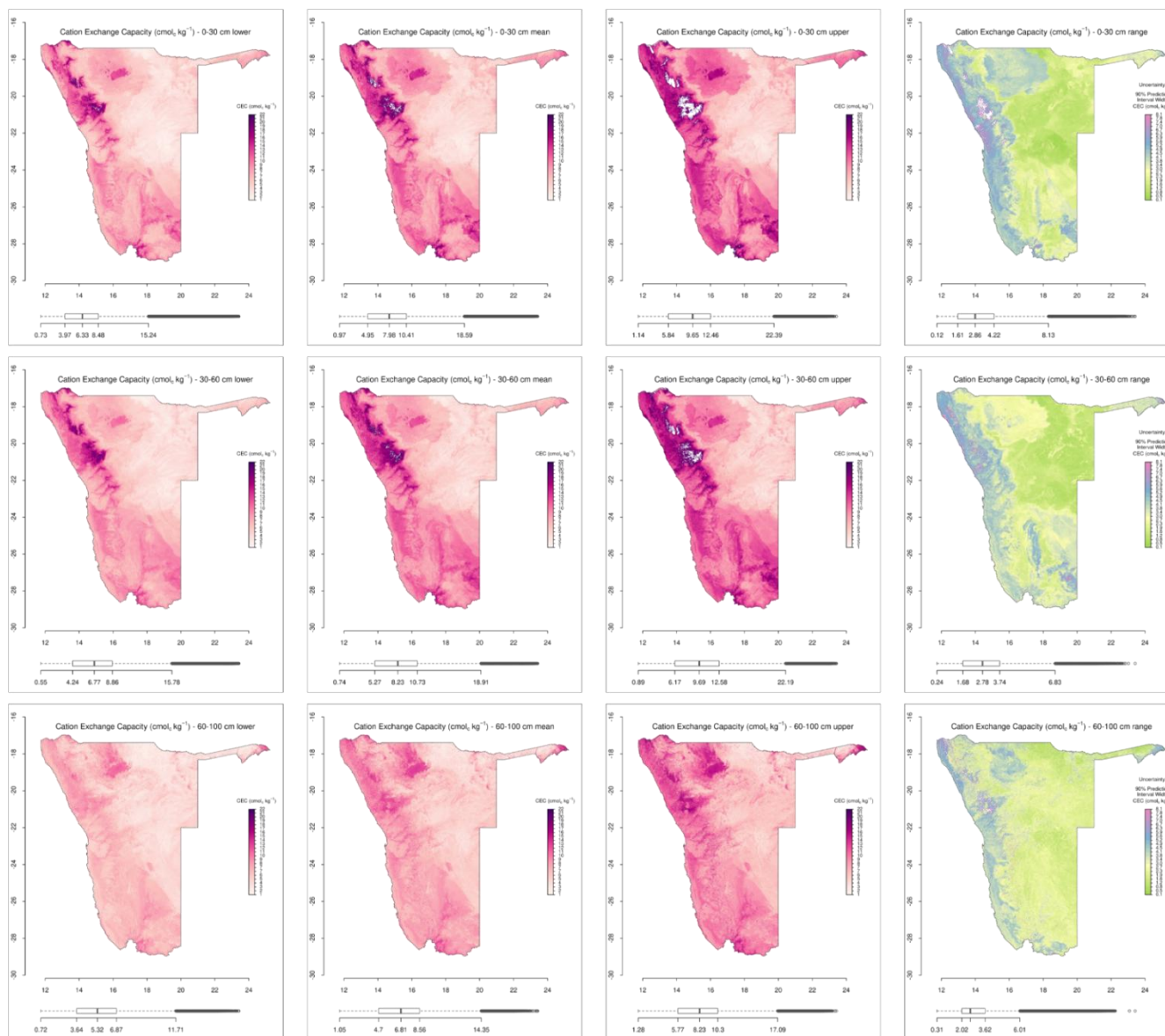
Base saturation



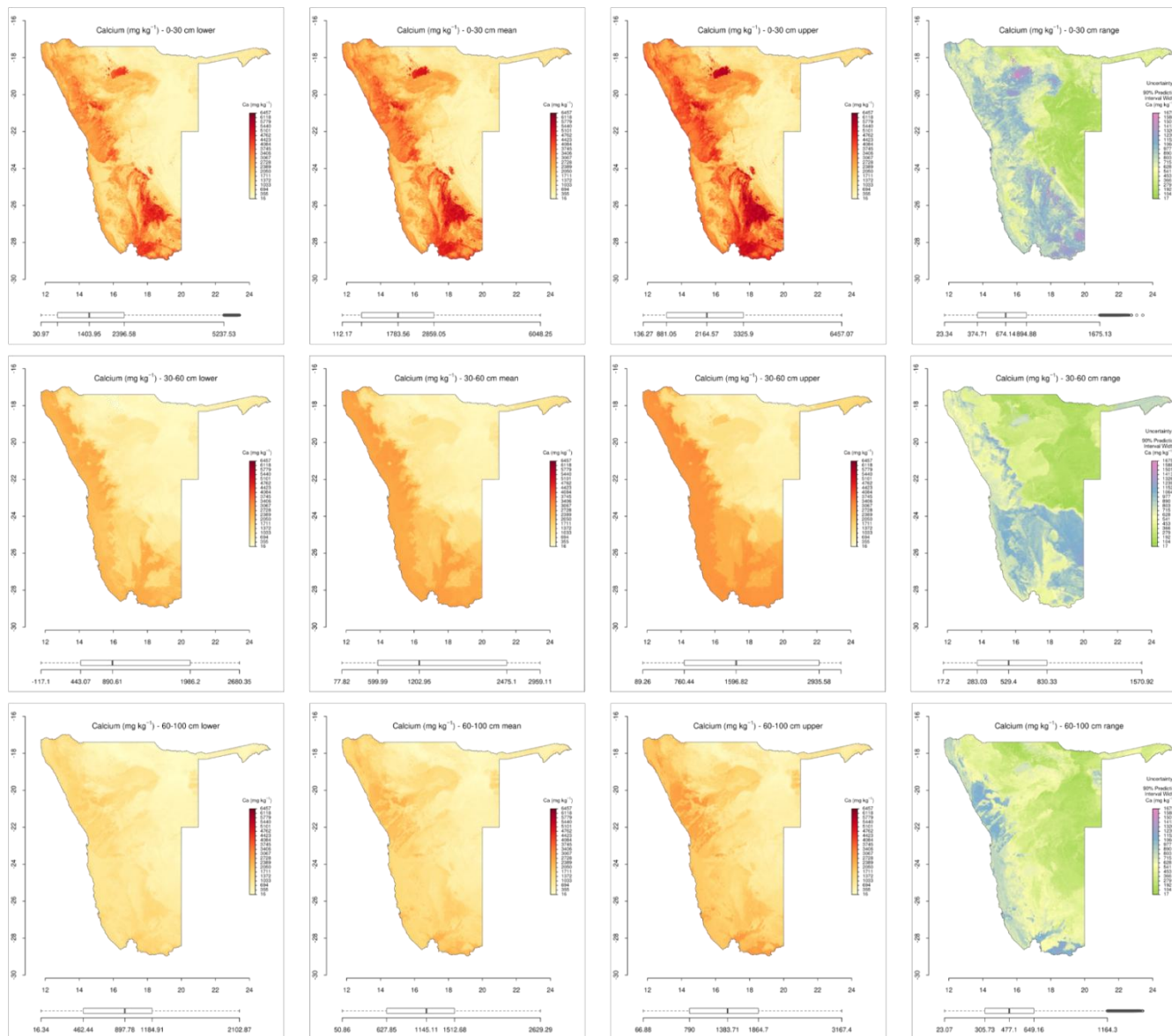
pH (H₂O)



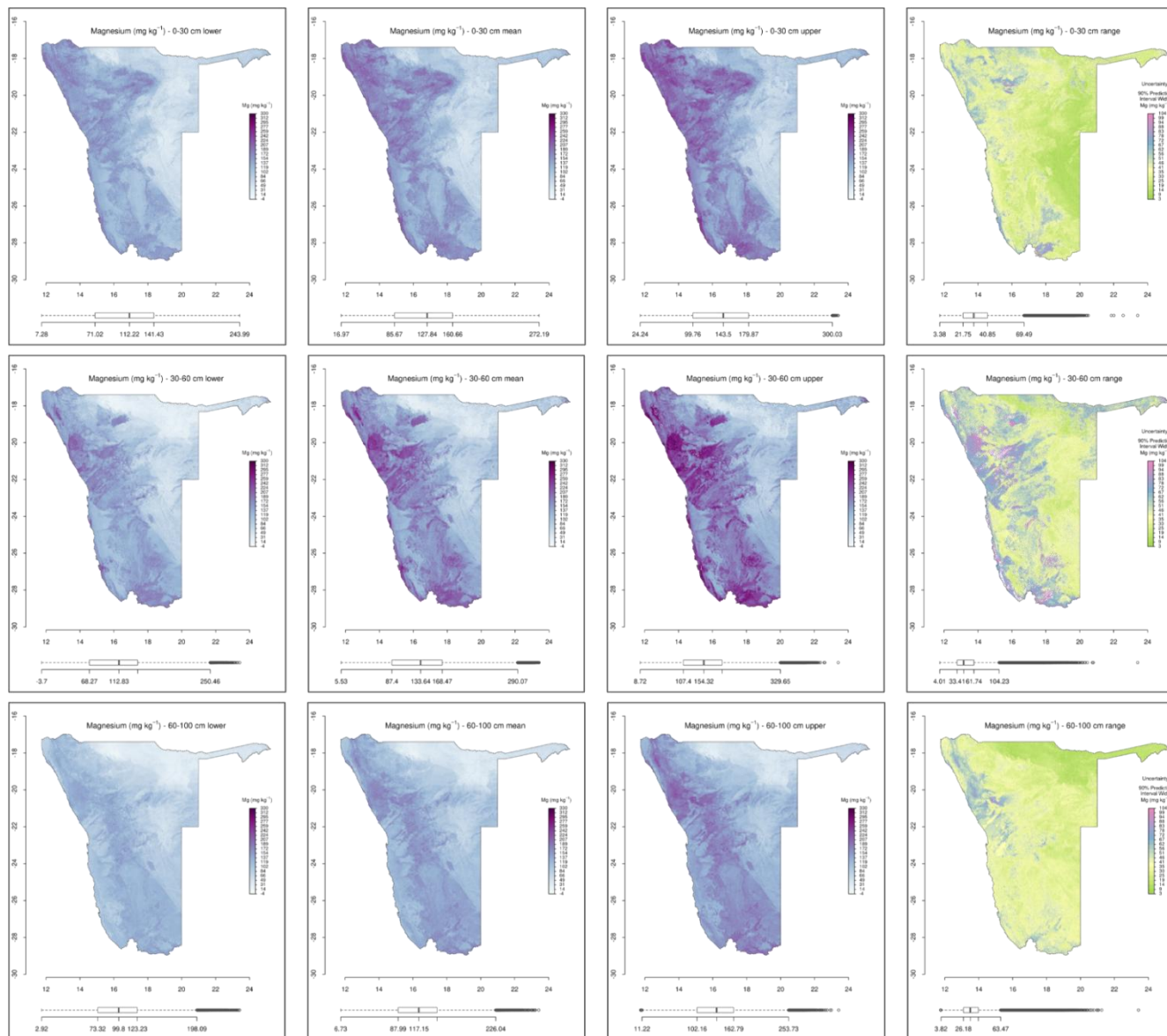
Cation exchange capacity



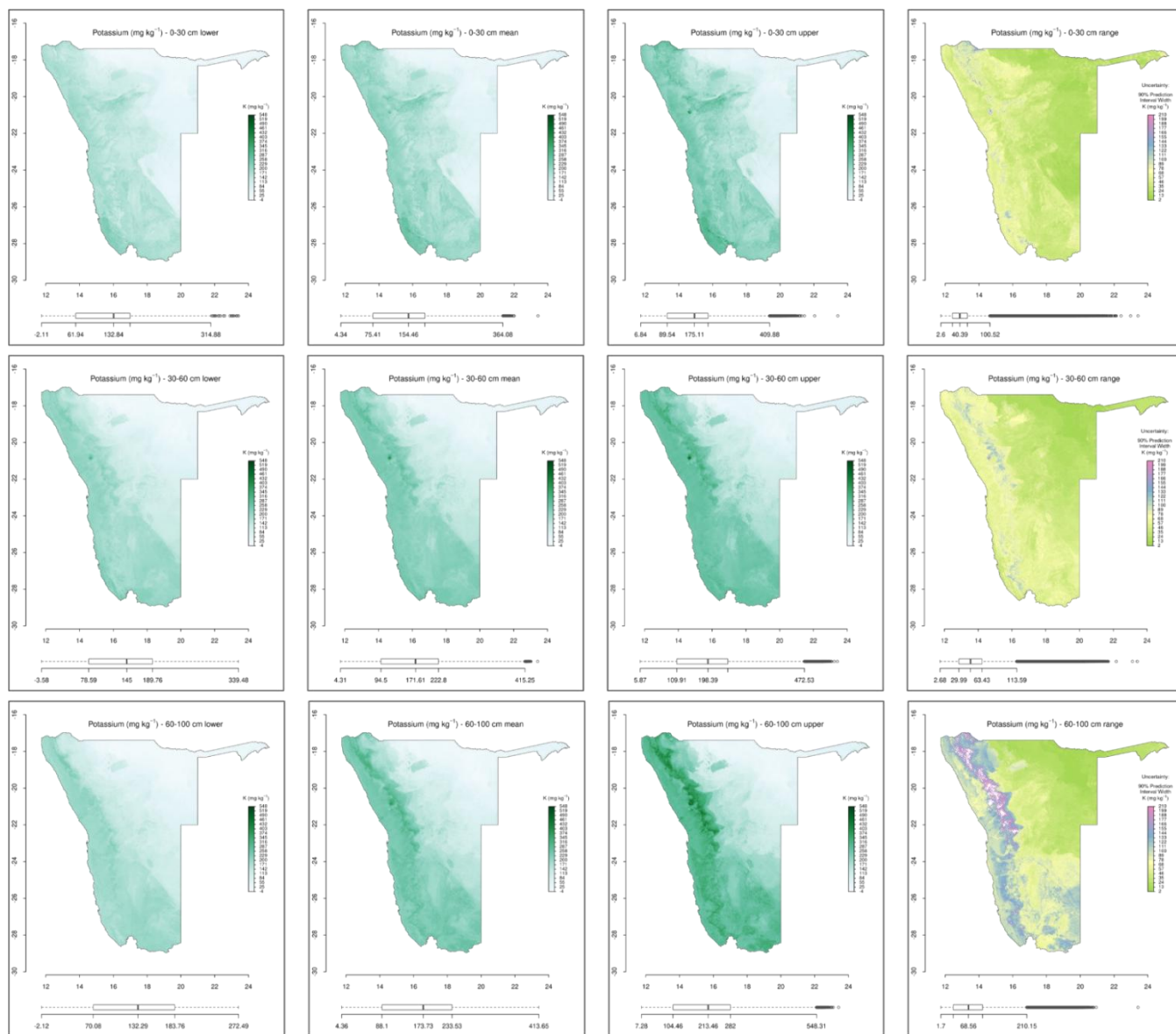
Extractable calcium



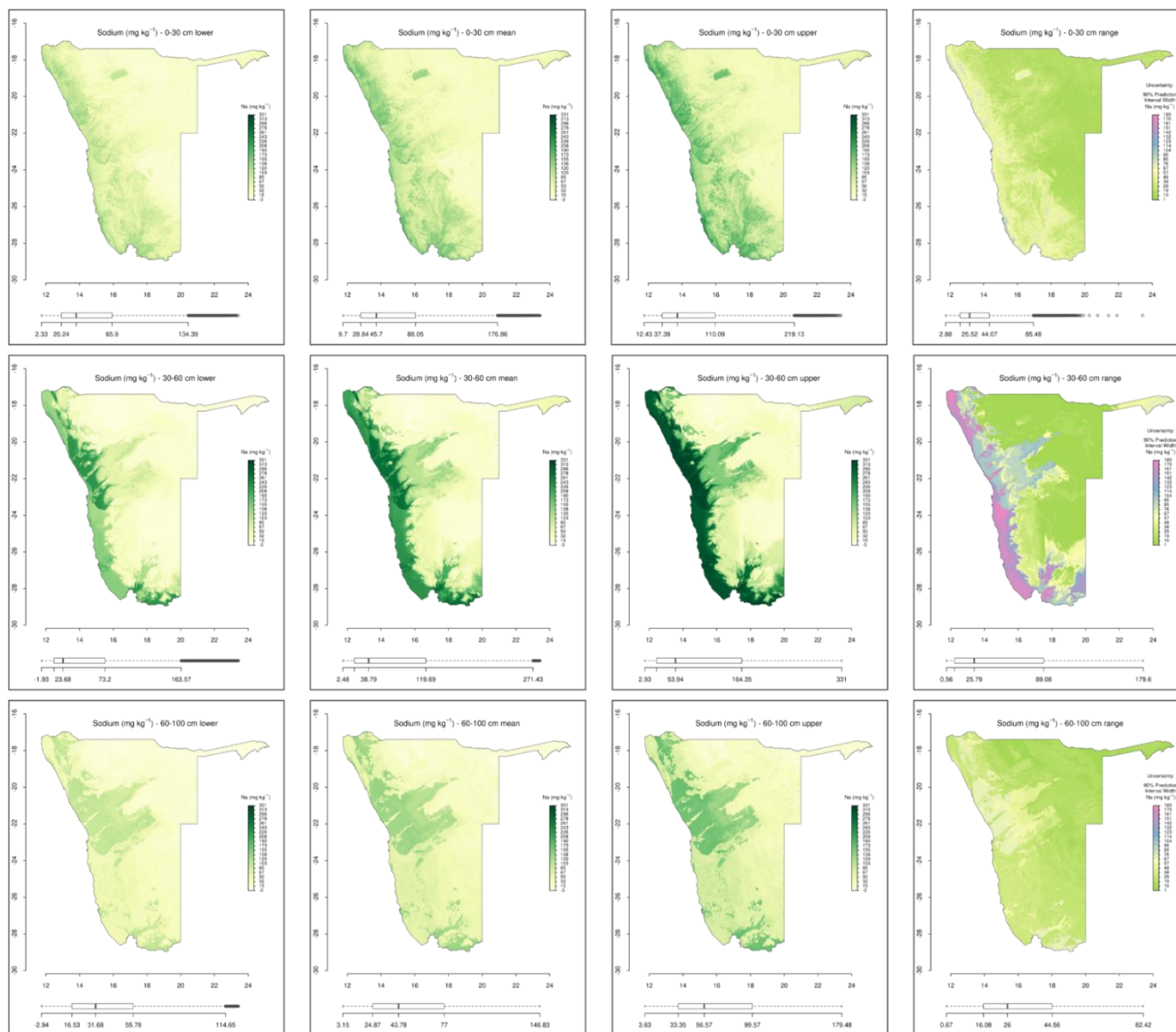
Extractable magnesium



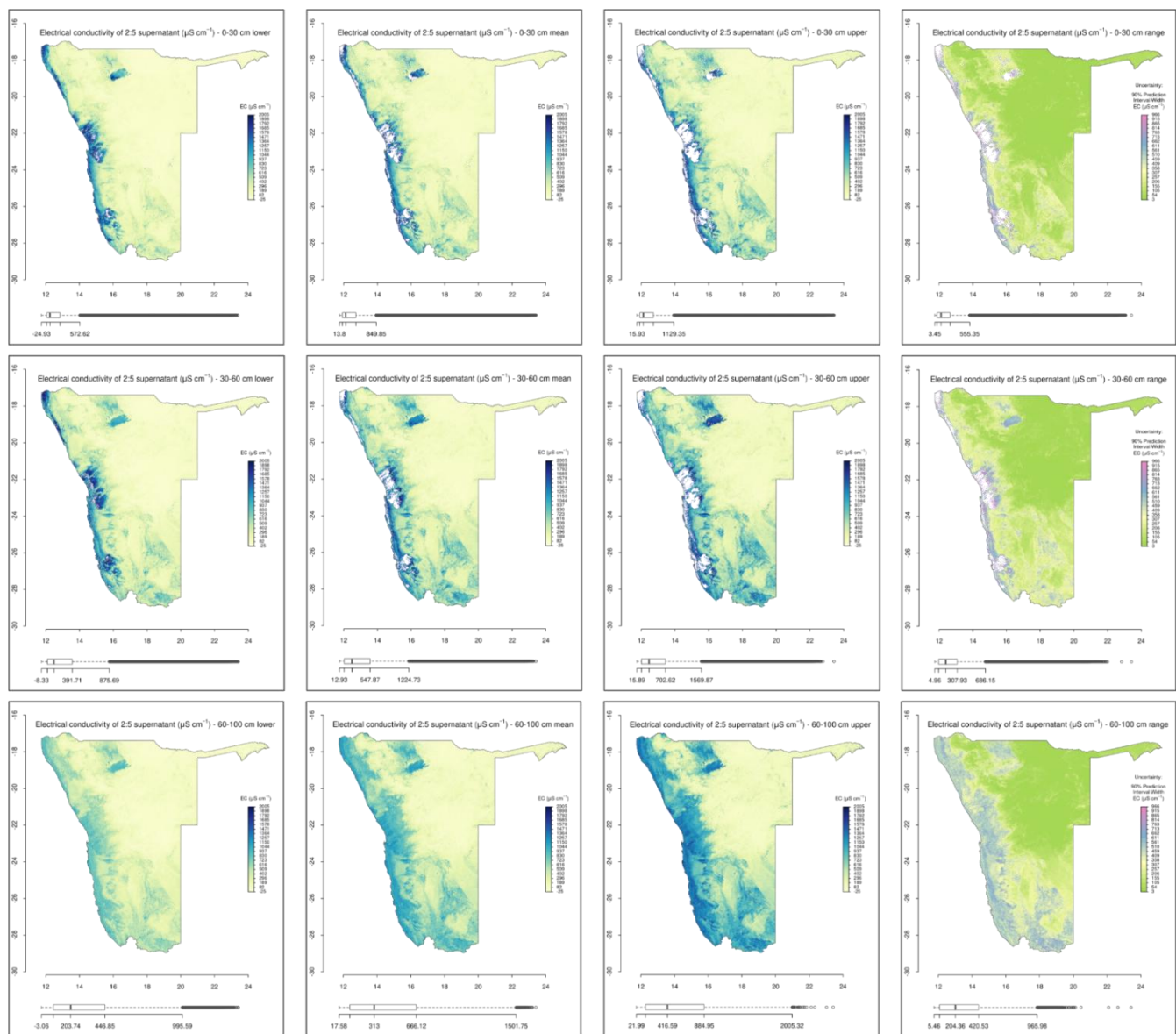
Extractable
potassium



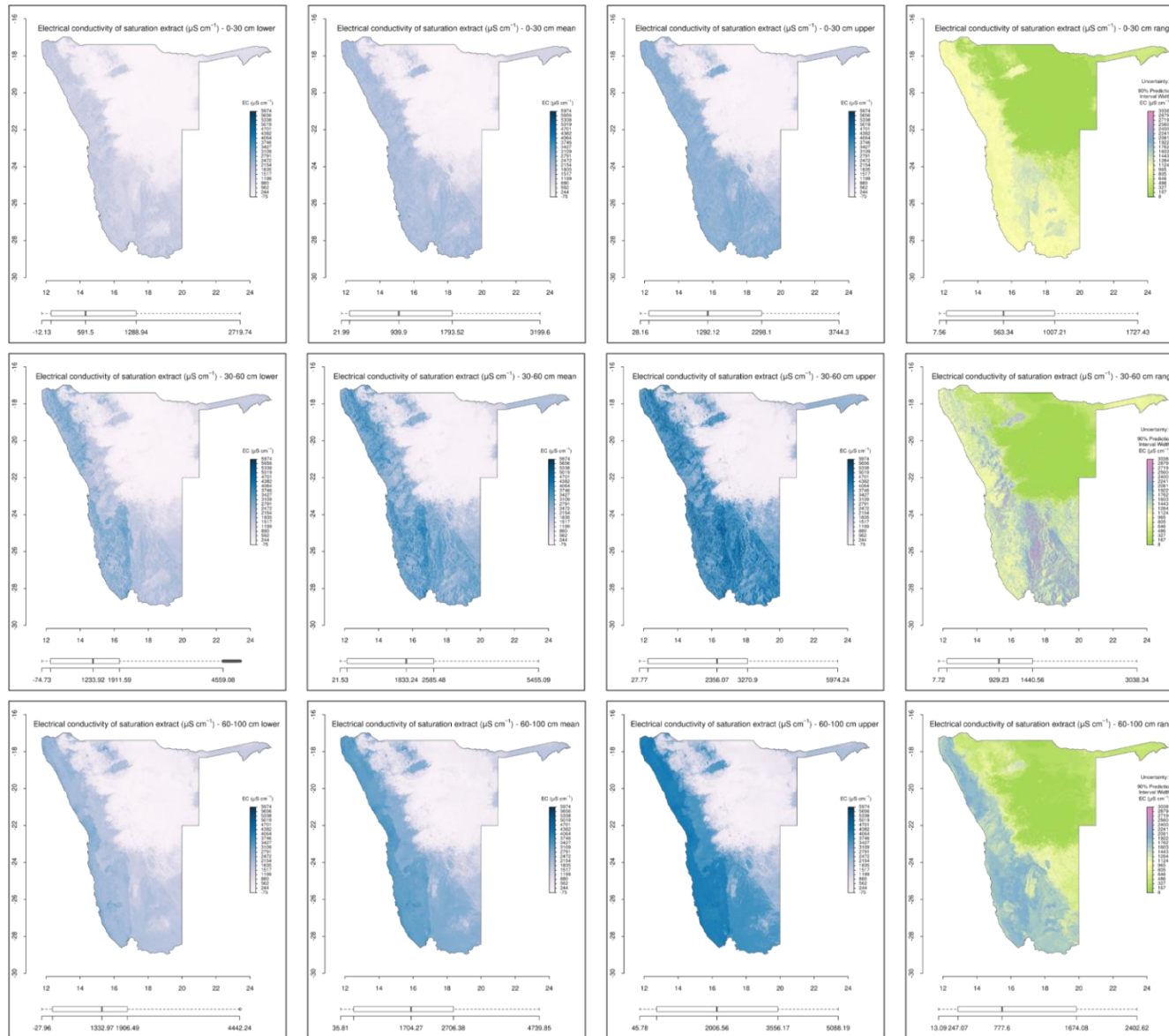
Extractable
sodium



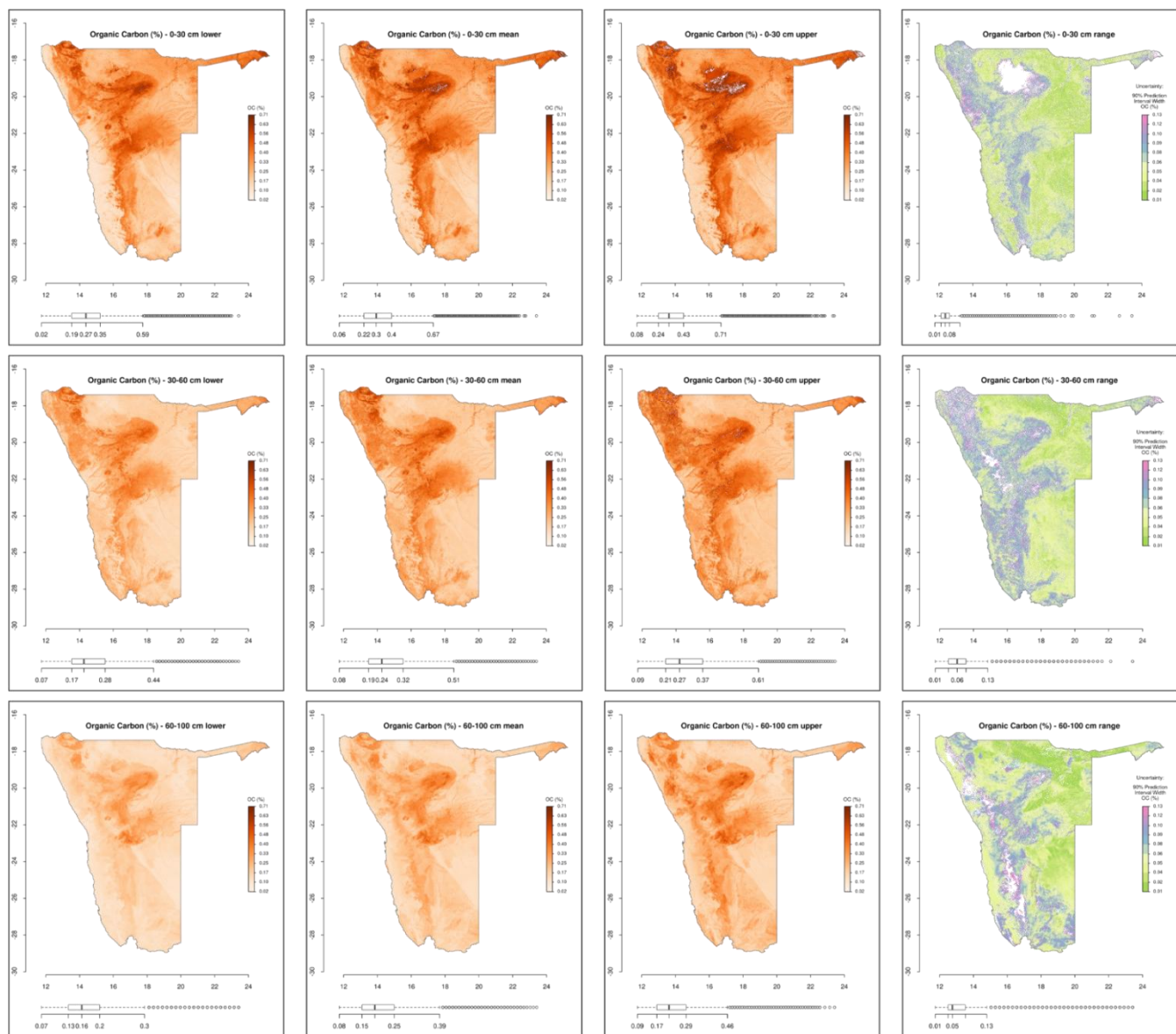
Electrical conductivity
(1:2.5)



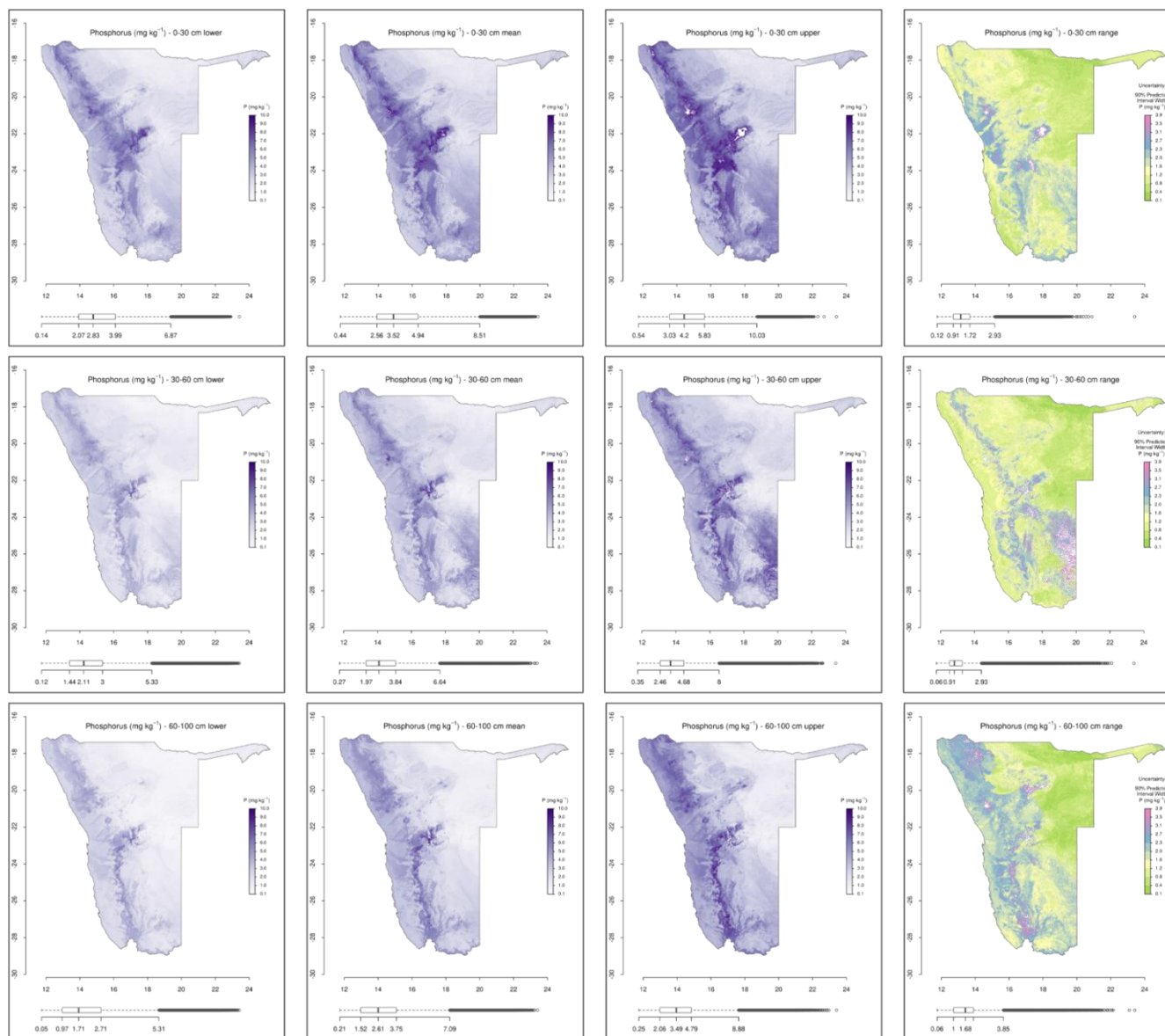
Electrical
conductivity
(saturated paste
extract)



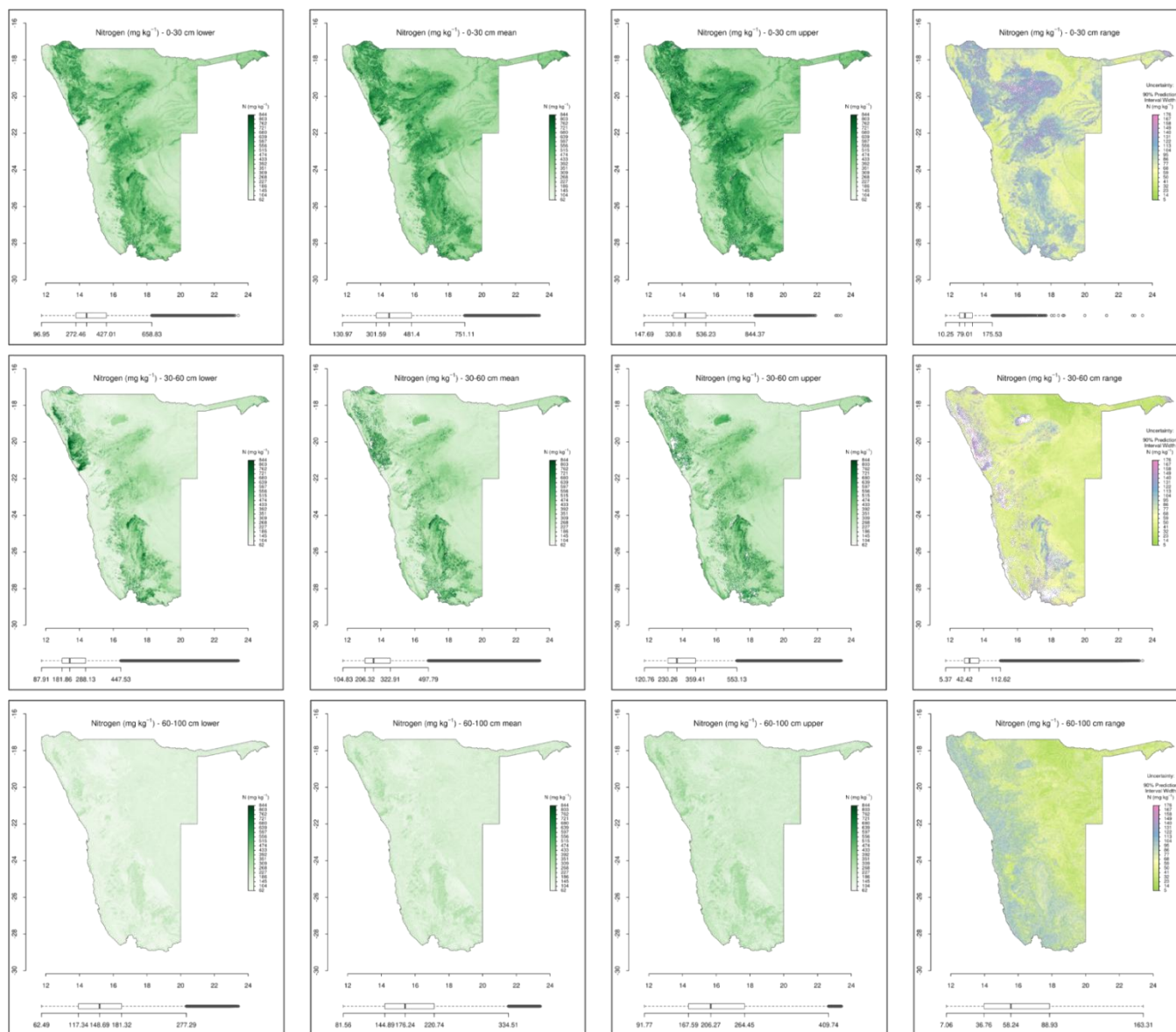
Organic carbon



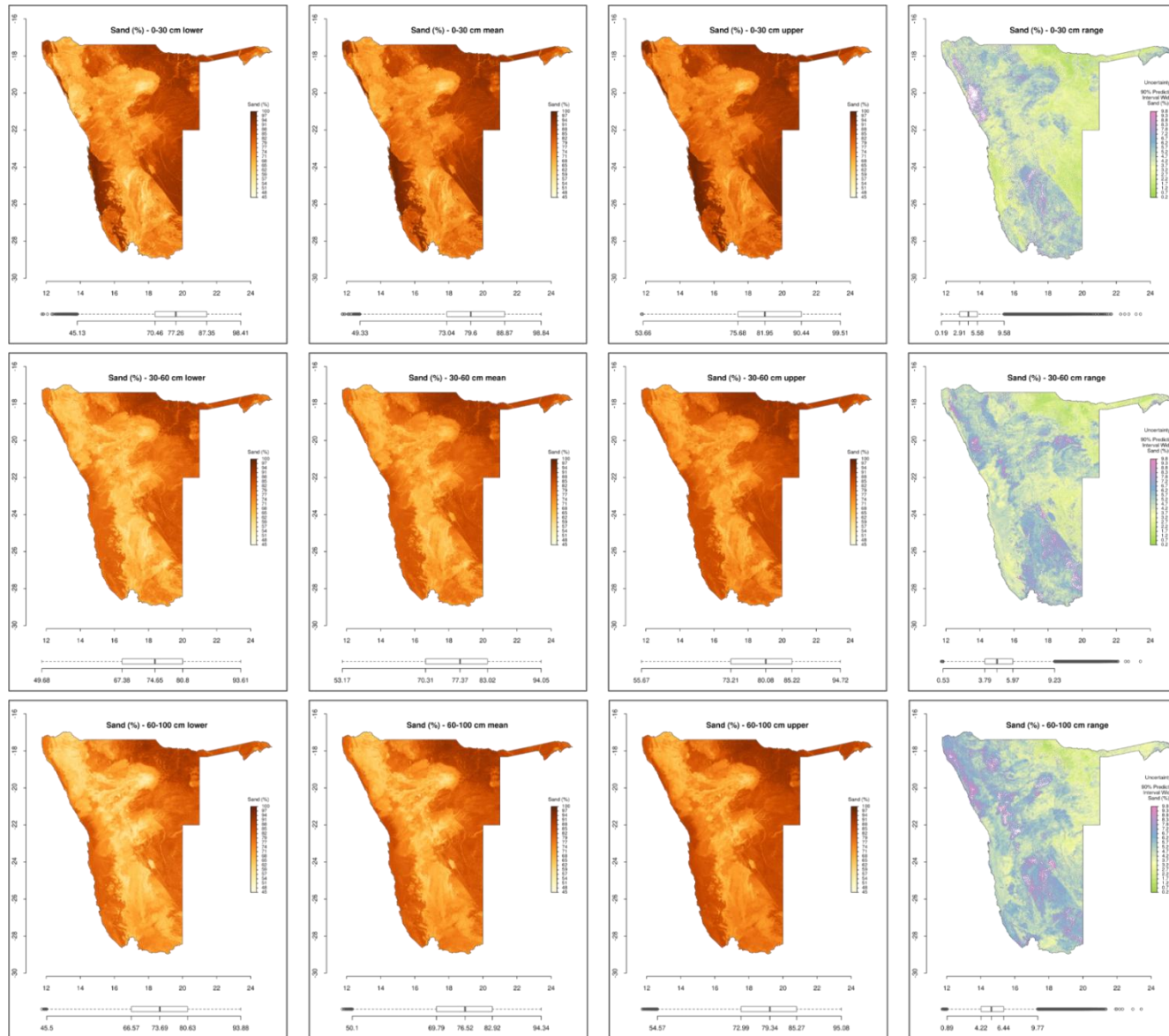
Plant-available
phosphorus



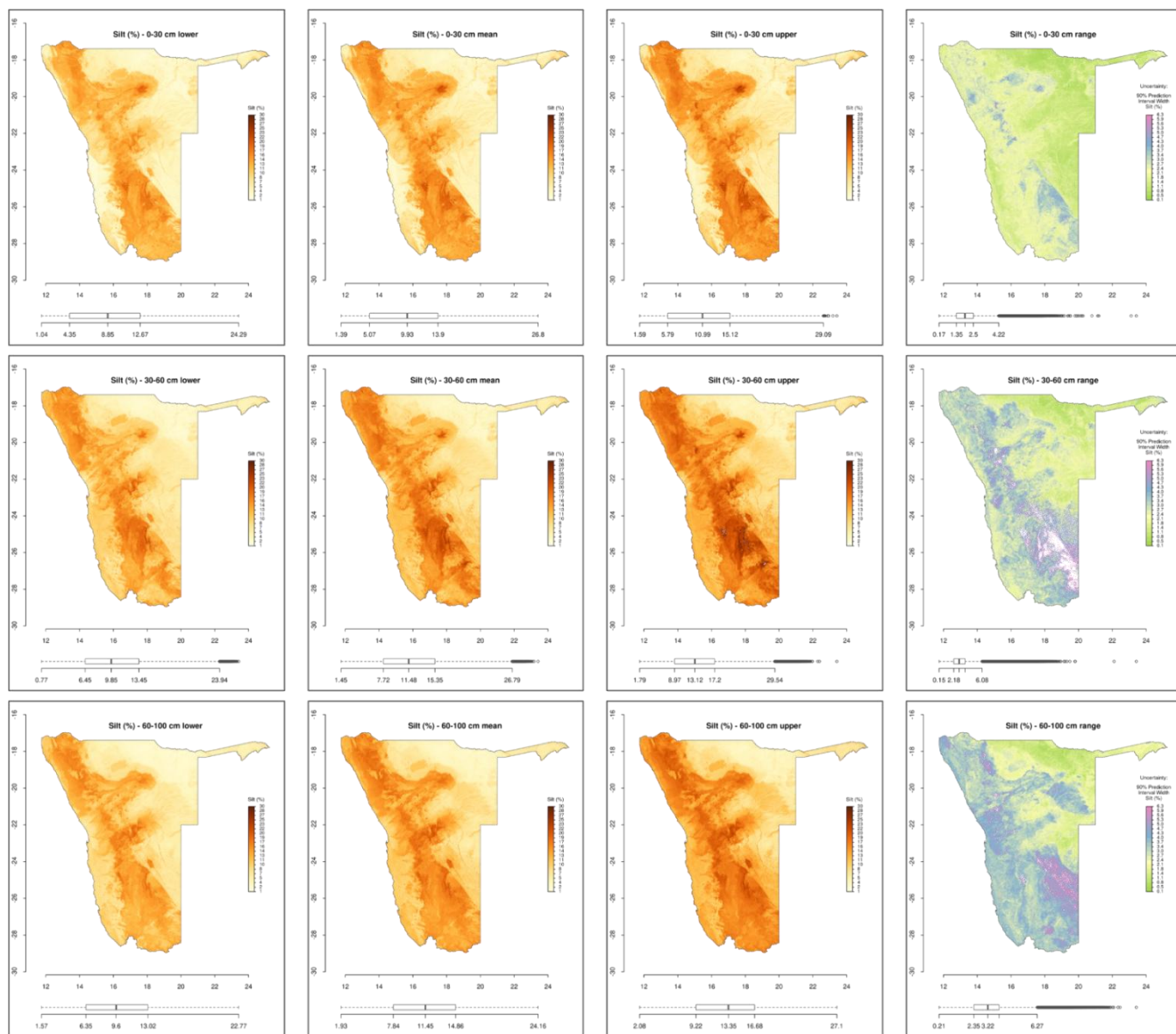
Total nitrogen



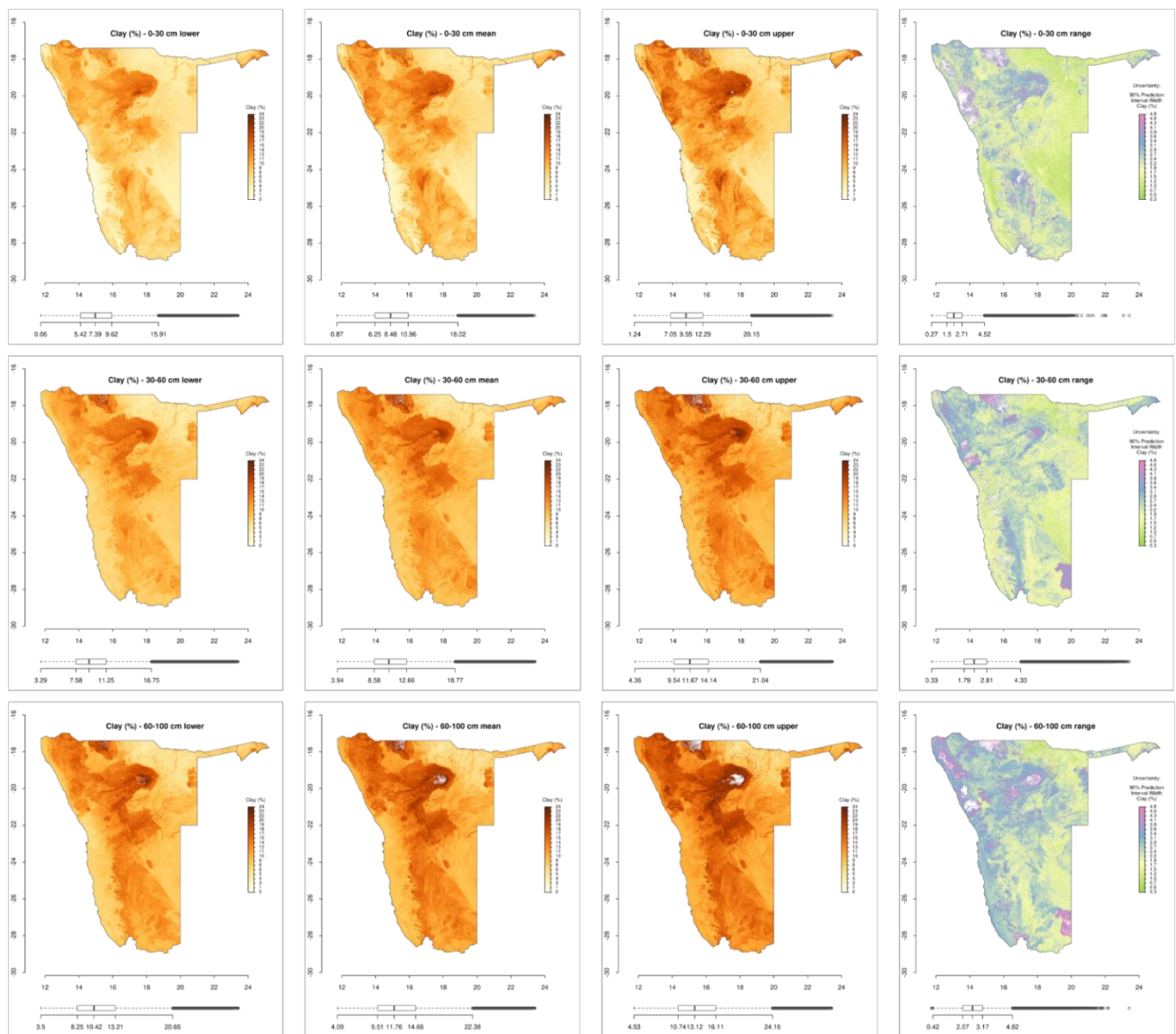
Sand



Silt

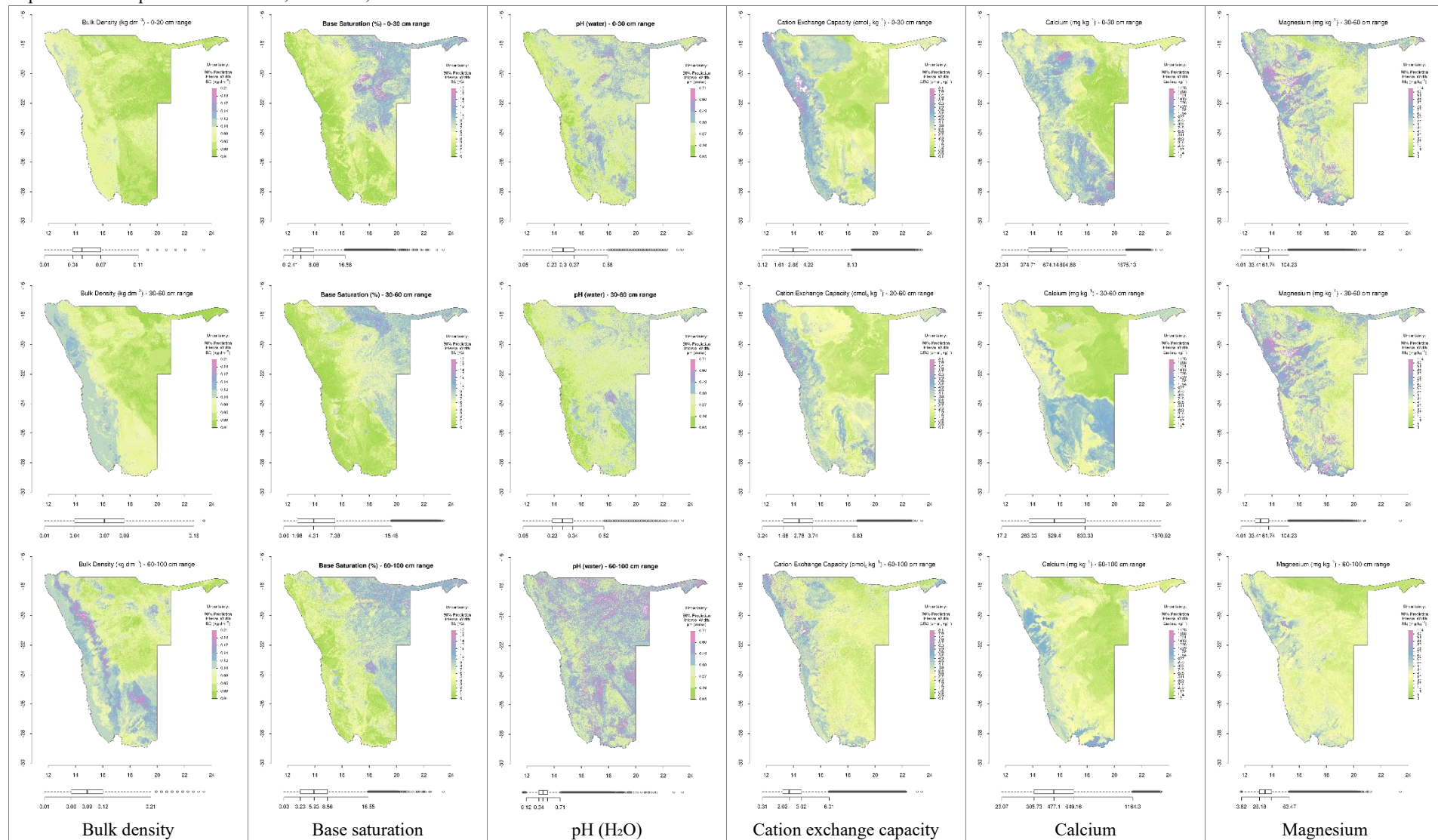


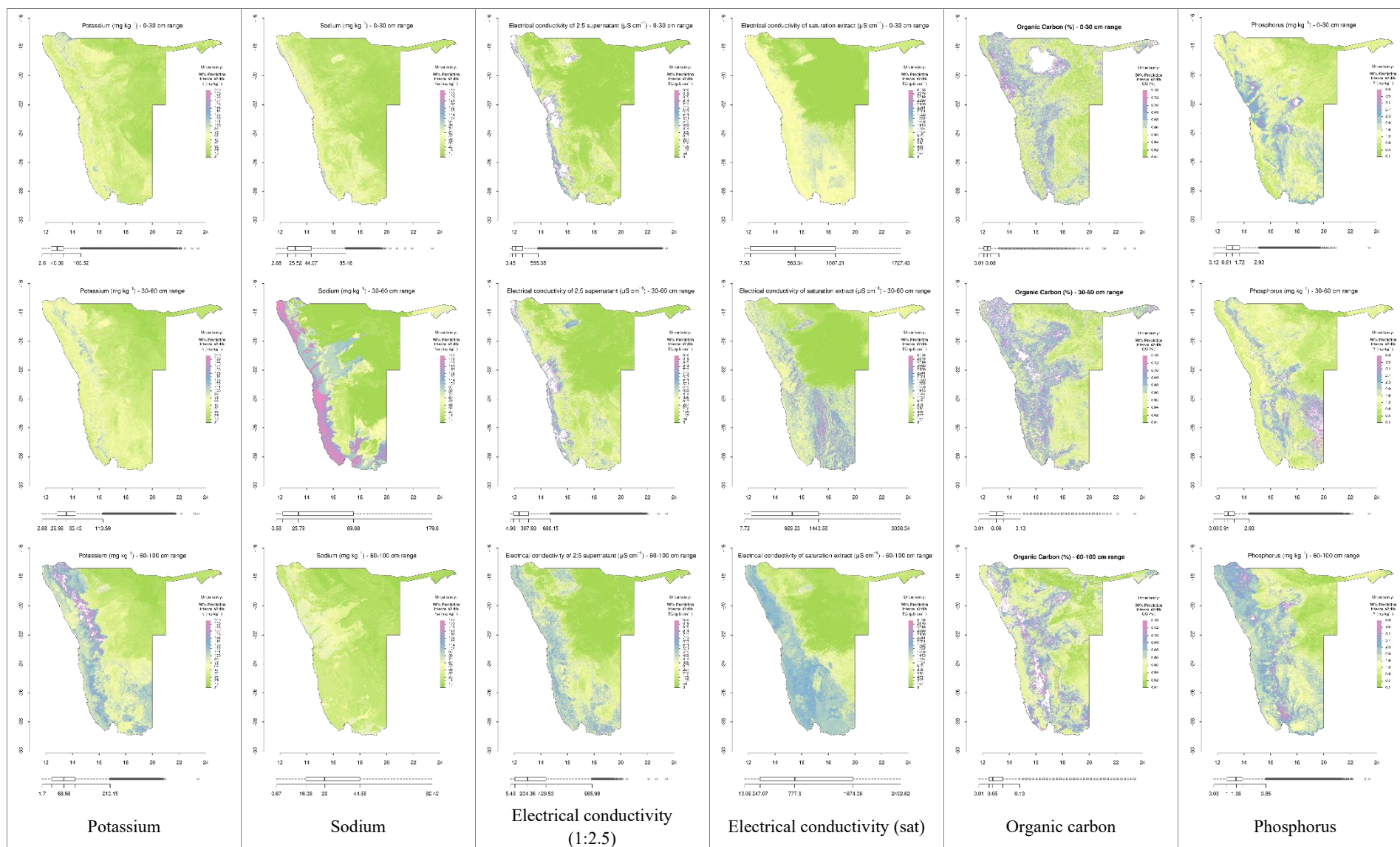
Clay

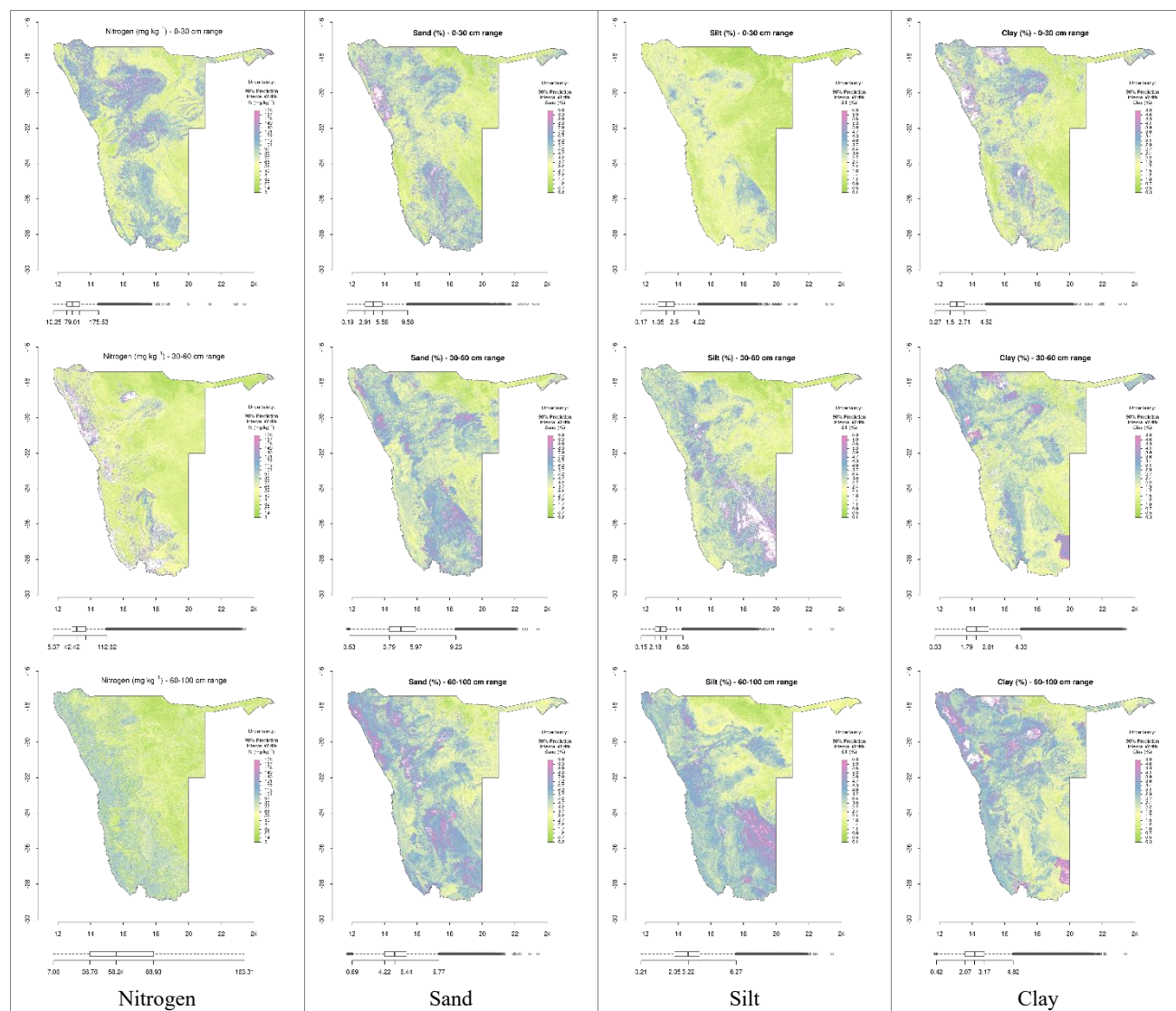


Q. PI90 maps for all soil properties and depth intervals

Top to bottom: depth classes 0–30 cm, 30–60 cm, 60–100 cm.

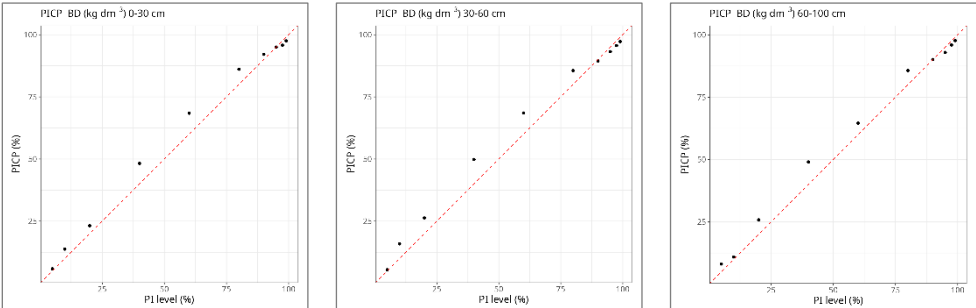




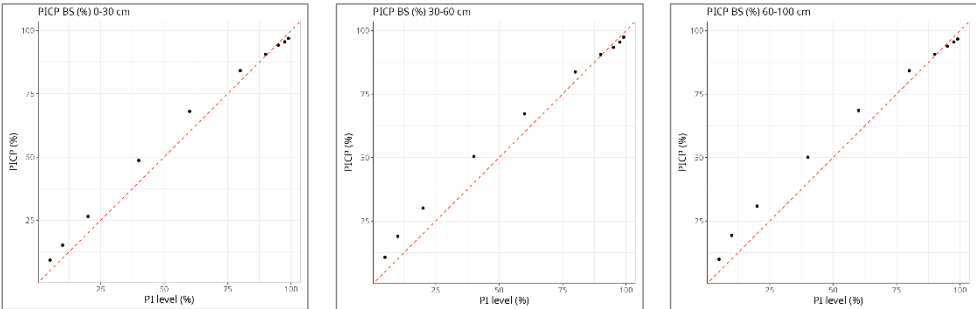


R. Prediction Interval Coverage Probability (PICP) plots for all soil properties and depth intervals.

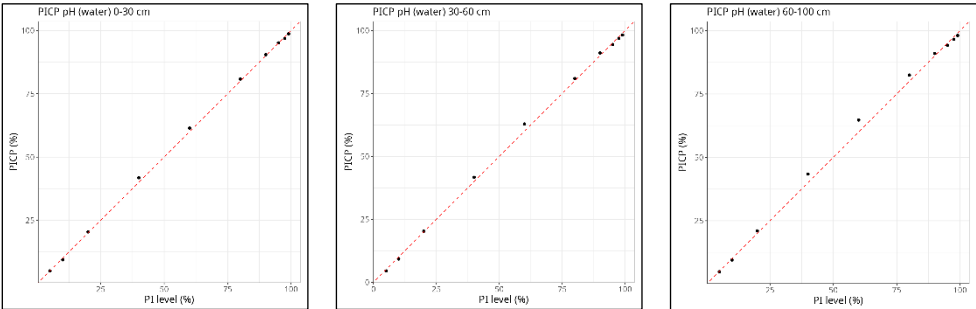
Bulk density



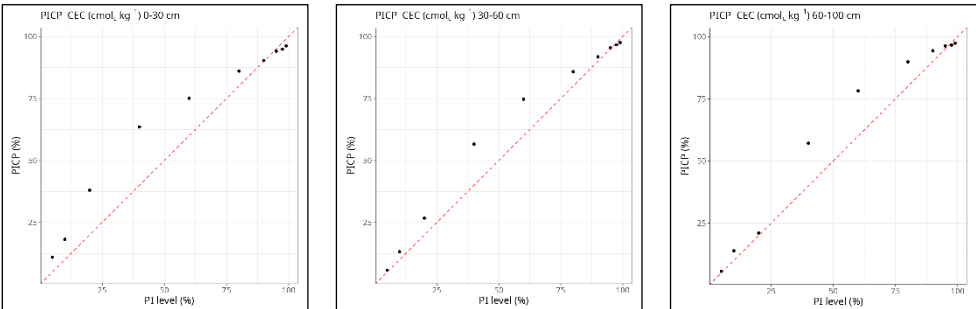
Base saturation



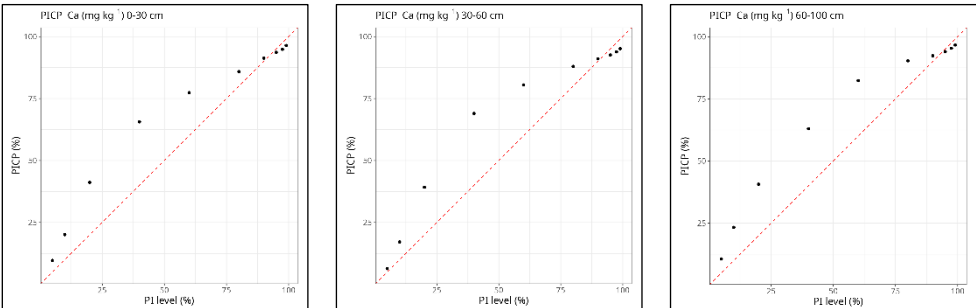
pH (H₂O)



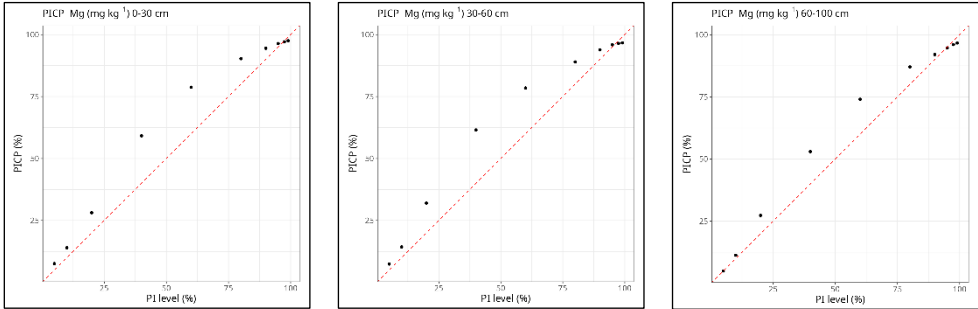
Cation exchange capacity



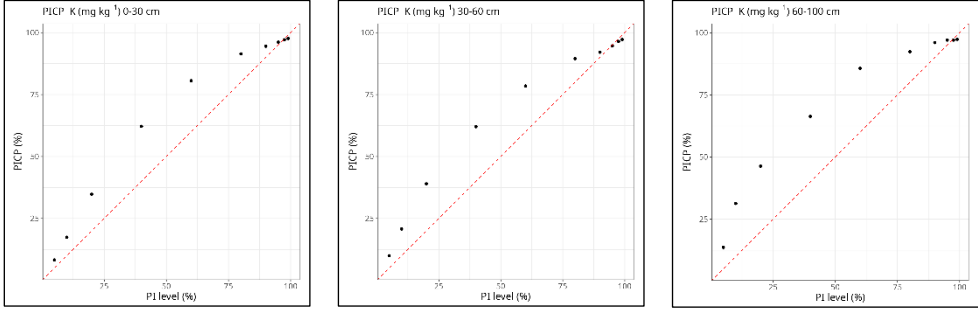
Calcium



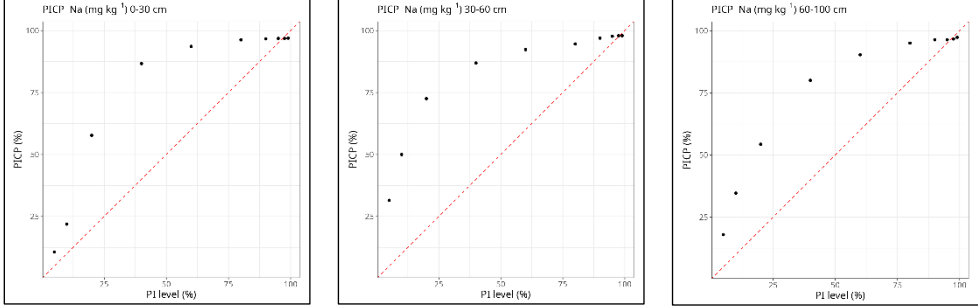
Magnesium



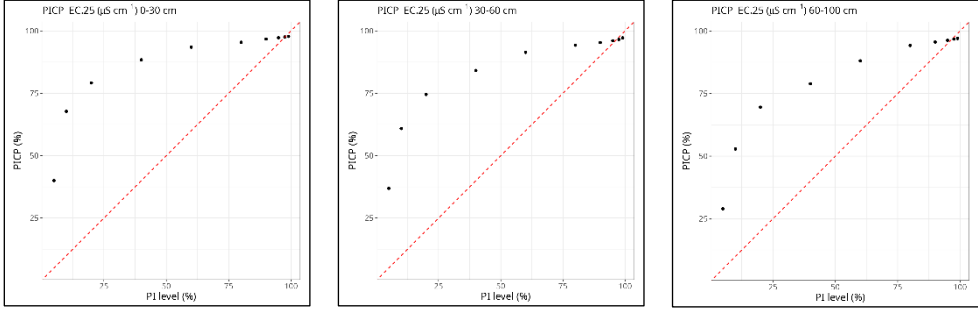
Potassium



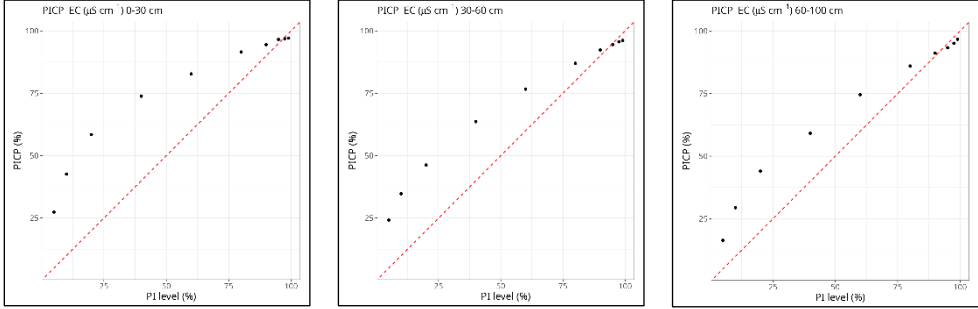
Sodium



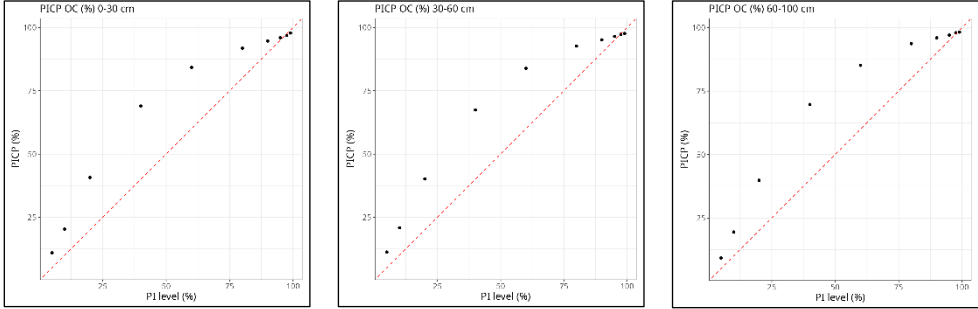
Electrical conductivity (1:2.5)



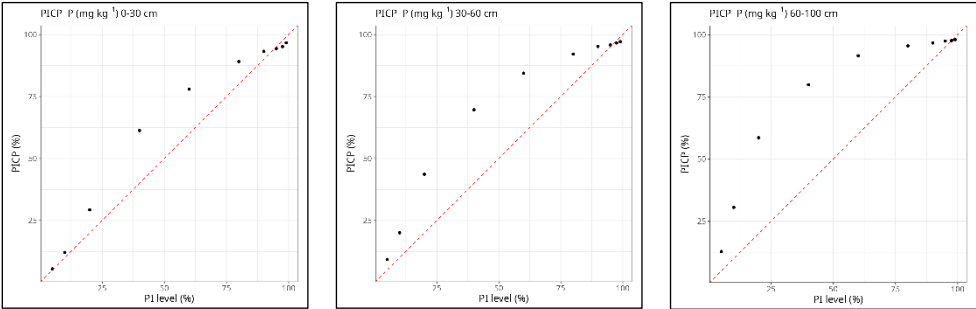
Electrical conductivity (sat)



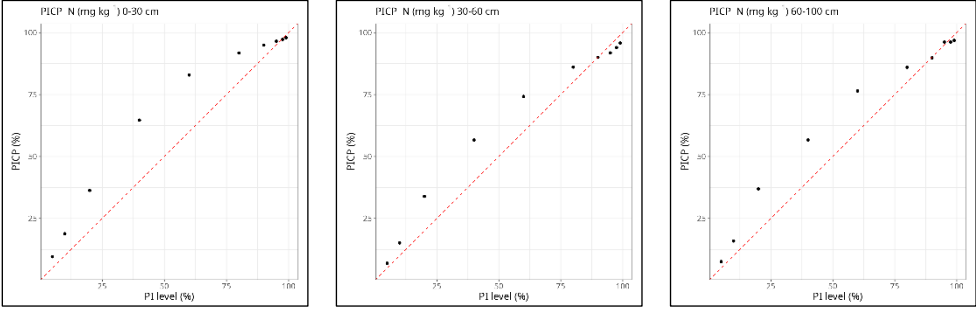
Organic carbon



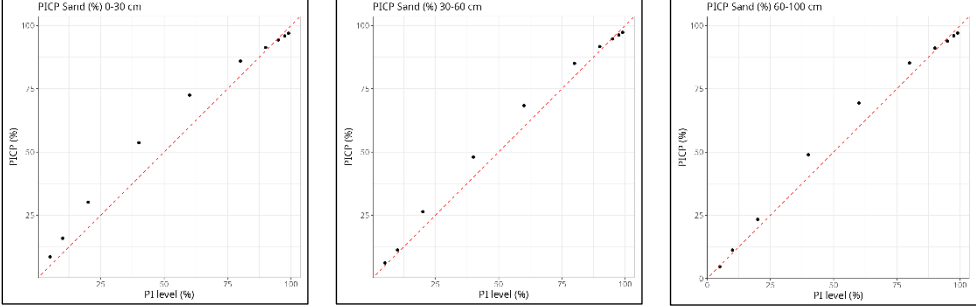
Phosphorus



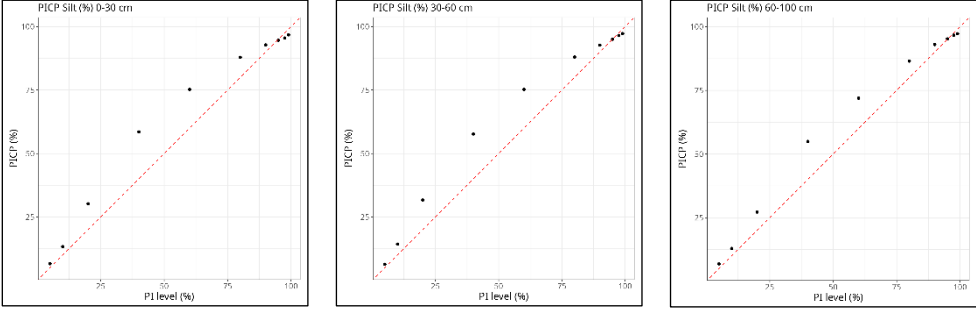
Nitrogen



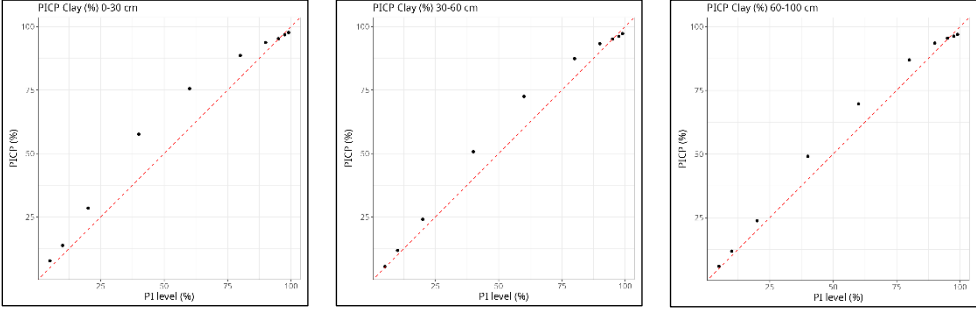
Sand



Silt

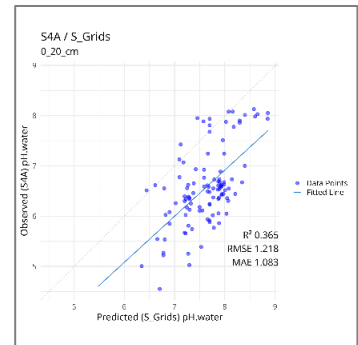
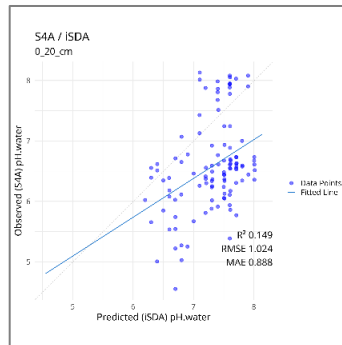
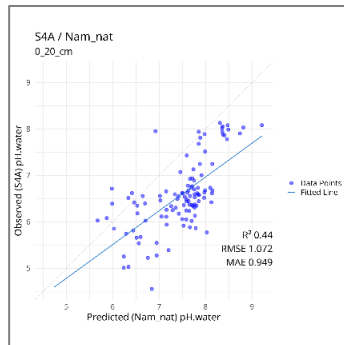


Clay

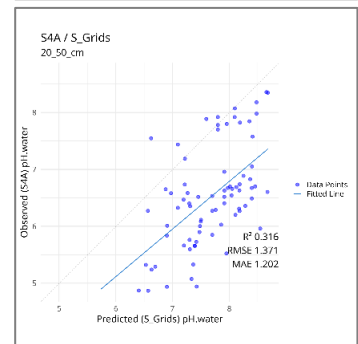
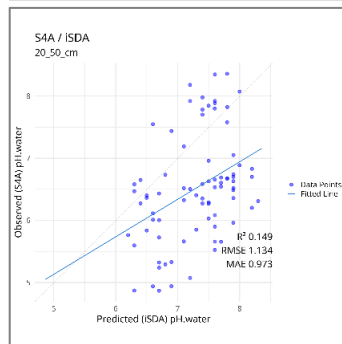
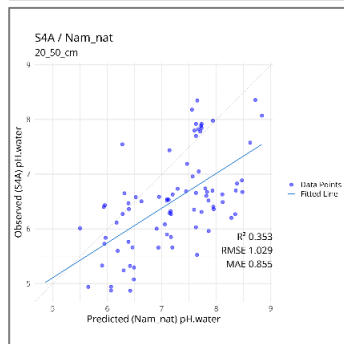


S. Scatterplots with validation statistics from external validation comparing predictions from this study, iSDA, and SoilGrids250m with independent Soils4Africa observations.

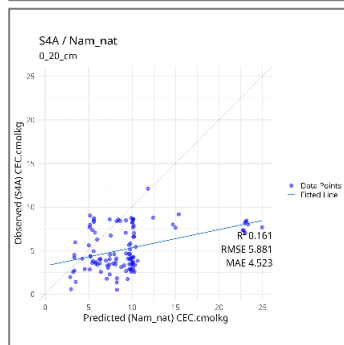
pH (H₂O)
0–20 cm



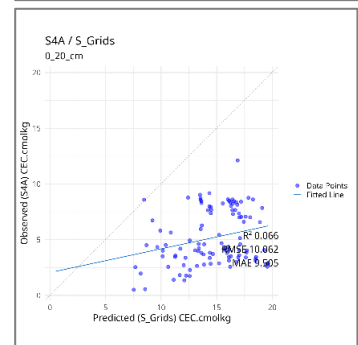
pH (H₂O)
20–50 cm



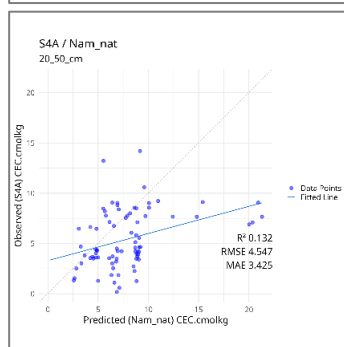
CEC
(cmol(+) kg⁻¹)
0–20 cm



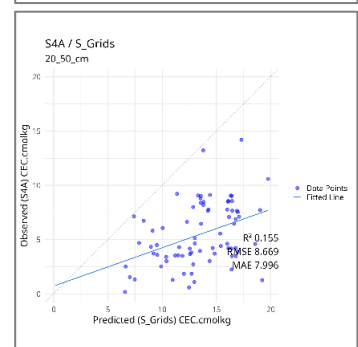
n/a



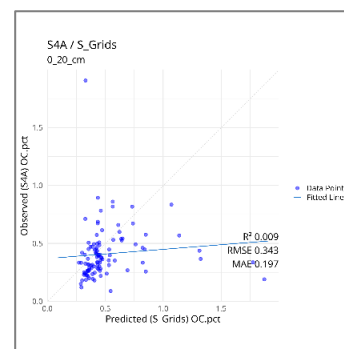
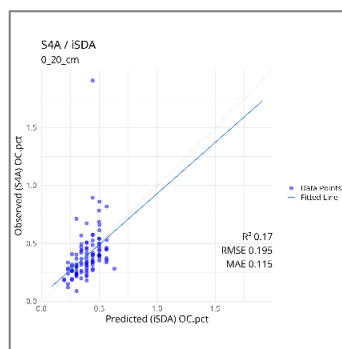
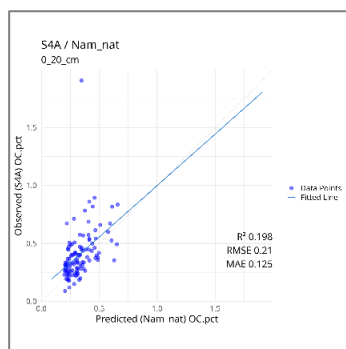
CEC
(cmol(+) kg⁻¹)
20–50 cm



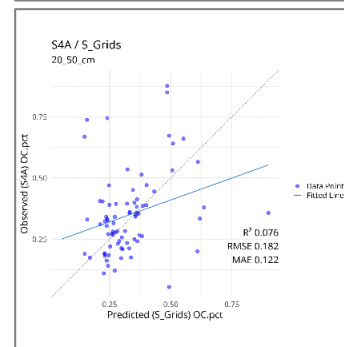
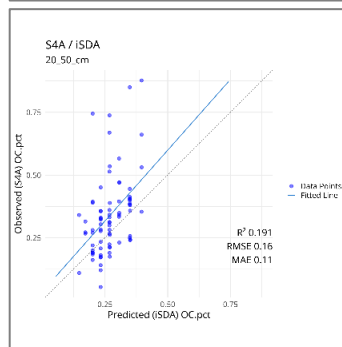
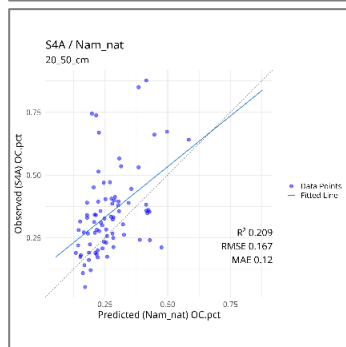
n/a



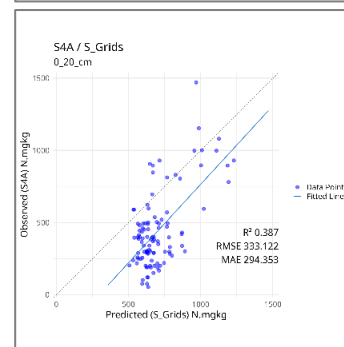
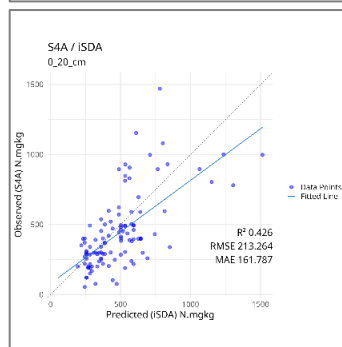
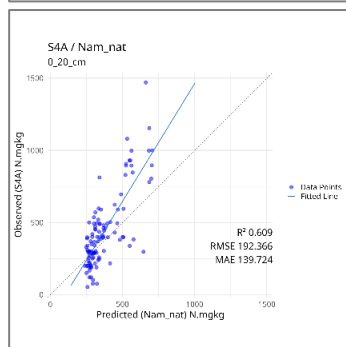
OC
(%)
0–20 cm



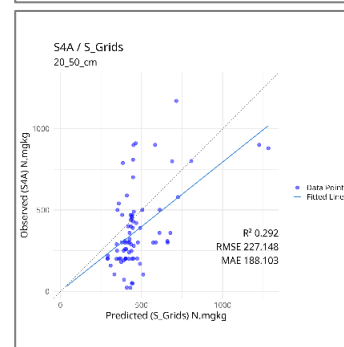
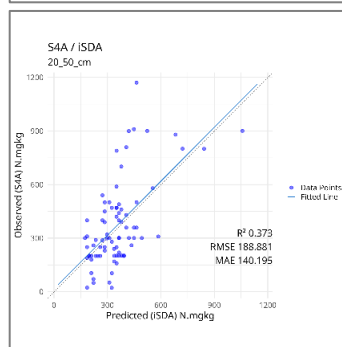
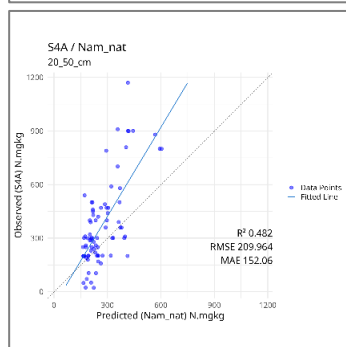
OC
(%)
20–50 cm



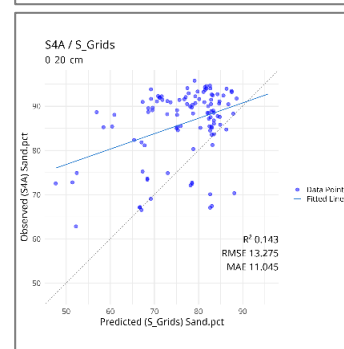
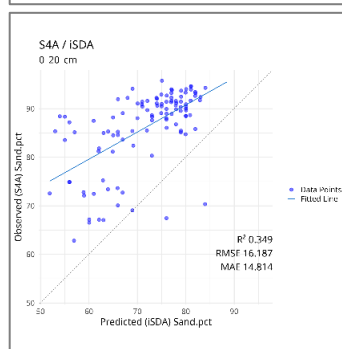
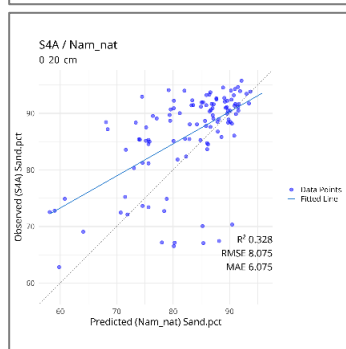
N
(mg kg⁻¹)
0–20 cm



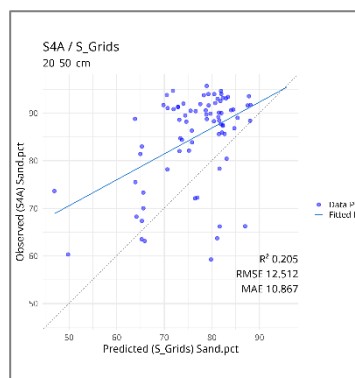
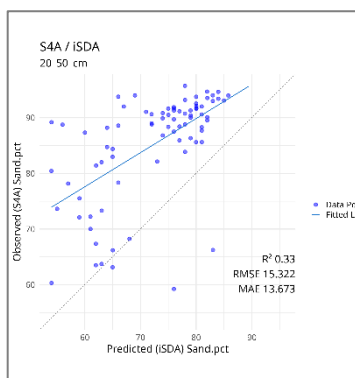
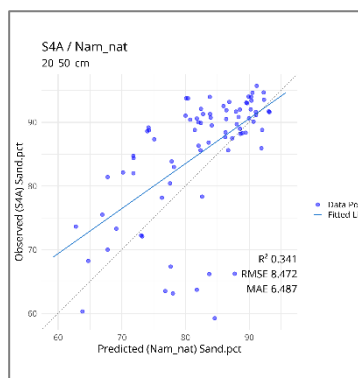
N
(mg kg⁻¹)
20–50 cm



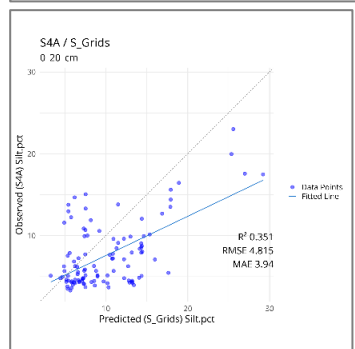
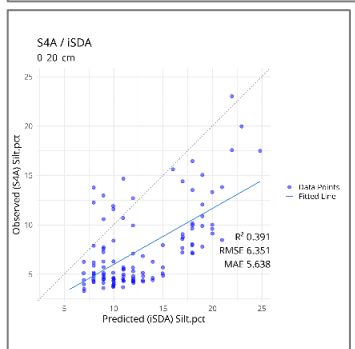
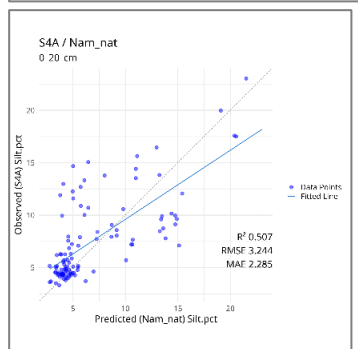
Sand
(%)
0–20 cm



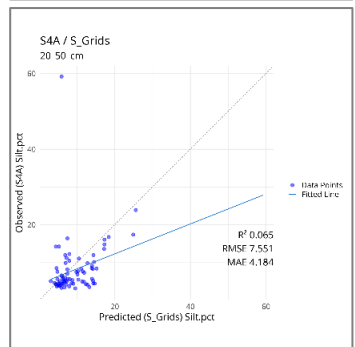
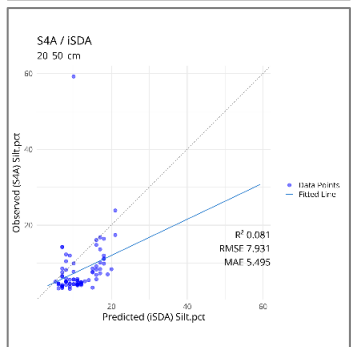
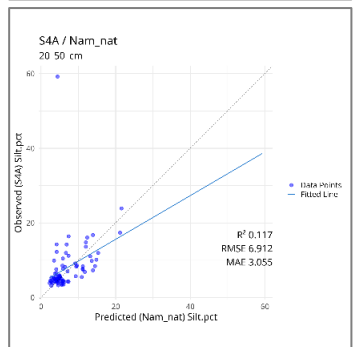
Sand
(%)
20–50 cm



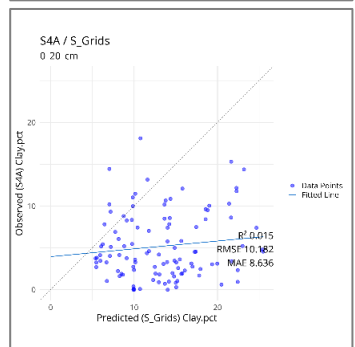
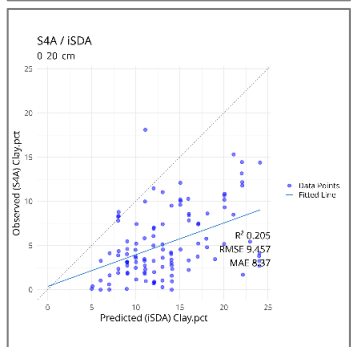
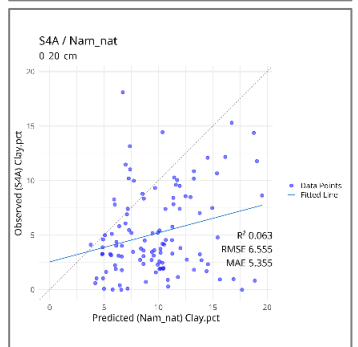
Silt
(%)
0–20 cm



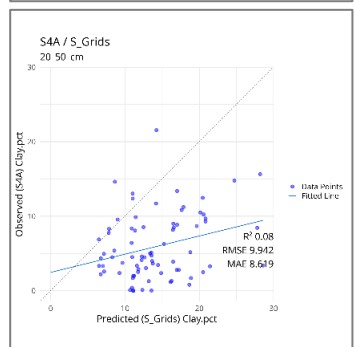
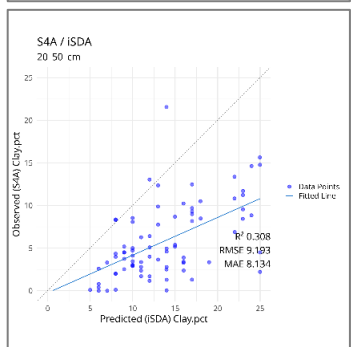
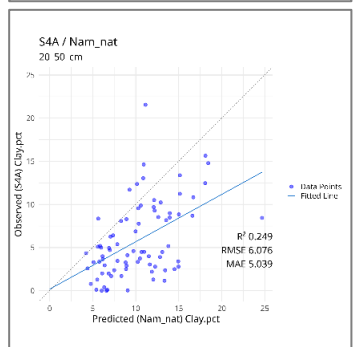
Silt
(%)
20–50 cm



Clay
(%)
0–20 cm



Clay
(%)
20–50 cm



T. Summary of model performance metrics reported in selected digital soil mapping studies.

Bulk density					
R ²	RMSE (kg dm ⁻³)	CCC	Source	Study Area	Notes
0.38, 0.19, 0.31	0.13, 0.12, 0.12		<i>This study</i>	Namibia	RF
0.318	0.23		Hengl et al., 2014	Global	SoilGrids1km; multiple linear regression
	0.14		Hengl et al., 2015	Africa	AfSoilGrids250m; RF
0.76	0.17		Hengl et al., 2017a	Global	SoilGrids250m 2016 update; Ensemble: RF, gradient boosting and/or multinomial logistic regression
0.82	0.13	0.901	Hengl et al., 2021	Africa	Ensemble ML models*
	0.13–0.15, 0.13, 0.13		Liu et al., 2022	China	Quantile regression forest
0.68–0.70, 0.67, 0.61	0.22–0.23, 0.24, 0.27		Padarian et al., 2017	Chile	CART; bootstrapping
	0.19		Poggio et al., 2021	Global	SoilGrids 2.0; Quantile regression forest
0.42	0.20		Ramcharan et al., 2017	Conterminous USA	RF, gradient boosting
pH (H ₂ O)					
R ²	RMSE	CCC	Source	Study Area	
0.51, 0.46, 0.48	0.87, 1.08, 1.13		<i>This study</i>	Namibia	
0.26	0.15		Bouslih et al., 2024	Morocco	RF
0.30	0.75	0.37	Dharumarajan et al., 2017	Anantapur District, Andhra Pradesh, India	RF
0.31	0.80	0.36	Ellur et al., 2024	Mandya district, India	RF
0.39–0.41	0.47–0.48		Guo et al., 2022	Fujian Province, China	RF, raw & with SVM-RFE feature selection
0.51	0.93		Hengl et al., 2014	Global	SoilGrids1km; multiple linear regression
	0.67		Hengl et al., 2015	Africa	AfSoilGrids250m; RF
0.83	0.50		Hengl et al., 2017a	Global	SoilGrids250m 2016 update; Ensemble: RF, gradient boosting and/or multinomial logistic regression
0.82	0.46	0.900	Hengl et al., 2021	Africa	Ensemble ML models*
0.52–0.59	0.66–0.72	0.67–0.75	Kasraie et al., 2024	British Columbia, Canada	RF; 0–30 cm
0.38	0.75	0.50	Kumaraperumal et al., 2022	Eastern Districts of Tamil Nadu, India	RF
	0.73–0.79, 0.73, 0.73		Liu et al., 2022b	China	Quantile regression forest
	0.55–0.76		Nussbaum et al., 2018	Switzerland	RF

0.58–0.66, 0.64	0.66, 0.58	0.57, 0.58		Padarian et al., 2017	Chile	CART; bootstrapping
	0.78			Poggio et al., 2021	Global	SoilGrids 2.0; Quantile regression forest
0.68	0.74			Ramcharan et al., 2017	Conterminous USA	RF, gradient boosting
0.66–0.73	0.37–0.38, 0.33	0.38, 0.82–0.84; 0.82; 0.84		Reddy et al., 2021	India	Modified 3D-RK, RF
0.36	0.65			Suleymanov et al., 2023	Saline steppe, Russia	RF; 0–20 cm
0.81	0.28			Vandana et al., 2024	Suryapet district, India	RF; 0–15 cm
Cation exchange capacity						
R²	RMSE (cmol(+) kg⁻¹)	CCC		Source	Study Area	
0.46, 0.49, 0.16	5.5, 5.1, 6.0			<i>This study</i>	Namibia	
0.33	12	0.36		Ellur et al., 2024	Mandya District, India	RF
0.29	11.4			Hengl et al., 2014	Global	SoilGrids1km; General Linear Models with log-link function
	7.92			Hengl et al., 2015	Africa	AfSoilGrids250m; RF
0.65	10.3			Hengl et al., 2017a	Global	SoilGrids250m 2016 update; Ensemble: RF, gradient boosting and/or multinomial logistic regression
	7.7–8.4, 7.5, 8.0			Liu et al., 2022b	China	Quantile regression forest
0.40	8.84	0.52		Kumaraperumal et al., 2022	Eastern Districts of Tamil Nadu, India	RF
	11.01			Poggio et al., 2021	Global	SoilGrids 2.0; Quantile regression forest
Calcium (extractable)						
R²	RMSE (mg kg⁻¹)	CCC		Source	Study Area	
0.46, 0.45, 0.32	1077, 832, 821			<i>This study</i>	Namibia	
0.69	1950			Hengl et al., 2017b	Africa	Ensemble RF & gradient boosting
0.84		0.913		Hengl et al., 2021	Africa	Mehlich3; Ensemble ML models*; ln-transformed
Magnesium (extractable)						
R²	RMSE (mg kg⁻¹)	CCC		Source	Study Area	
0.32, 0.27, 0.29	77, 102, 92			<i>This study</i>	Namibia	
0.78	241			Hengl et al., 2017b	Africa	Ensemble RF & gradient boosting
0.82		0.898		Hengl et al., 2021	Africa	Mehlich3; Ensemble ML models*; ln-transformed

Potassium (extractable)					
R ²	RMSE (mg kg ⁻¹)	CCC	Source	Study Area	
0.41, 0.41, 0.36	84, 90, 95		<i>This study</i>	Namibia	
0.64	201		Hengl et al., 2017b	Africa	Ensemble RF & gradient boosting
0.77		0.872	Hengl et al., 2021	Africa	Mehlich3; Ensemble ML models*; ln-transformed
Sodium (extractable)					
R ²	RMSE (mg kg ⁻¹)	CCC	Source	Study Area	
0.10, 0.14, 0.18	95, 110, 78				
0.61	452		Hengl et al., 2017b	Africa	Ensemble RF & gradient boosting
Electrical conductivity (1:2.5)					
R ²	RMSE (µS cm ⁻¹)	CCC	Source	Study Area	
0.37, 0.20, 0.10	541, 766, 750		<i>This study</i>	Namibia	
0.62	140	0.70	Dharumarajan et al., 2017	Anantapur District, Andhra Pradesh, India	RF; soil:water ratio unknown
0.18	223	0.20	Ellur et al., 2024	Mandya District, India	RF
0.73	134		Vandana et al., 2024	Suryapet District, India	RF; 0-15 cm
Organic carbon					
R ²	RMSE (%)	CCC	Source	Study Area	
0.32, 0.18, 0.16	0.24, 0.21, 0.22		<i>This study</i>	Namibia	
0.41–0.43, 0.29, 0.23	0.24–0.66, 0.56, 0.48		Adhikari et al., 2014	Denmark	Cubist
0.90	0.37		Bouslihim et al., 2024	Morocco [SOM]	RF; SOM, not SOC
0.23	0.20	0.38	Dharumarajan et al., 2017	Anantapur District, Andhra Pradesh, India	RF
0.34	0.43	0.40	Ellur et al., 2024	Mandya District, India	RF
0.58 ± 0.05	0.82 ± 0.03	0.78 ± 0.03	Emadi et al., 2020	Mazandaran Province, Iran	RF; 0–20 cm
0.229	1.3?		Hengl et al., 2014	Global	SoilGrids1km; General Linear Models with log-link function; dry combustion method
0.61	1.06?		Hengl et al., 2015	Africa	AfSoilGrids250m; RF
0.64	0.33		Hengl et al., 2017a	Global	SoilGrids250m 2016 update; Ensemble: RF, gradient boosting and/or multinomial logistic regression
0.79	0.37	0.883	Hengl et al., 2021	Africa	Ensemble ML models*; ln-transformed
0.24–0.30	3.26–3.40?	0.39–0.46	Kasraie et al., 2024	British Columbia, Canada	RF; 0–30 cm
0.42	0.38	0.549	Kumar et al., 2023	Upper Brahmaputra Valley, India	0–15 cm

0.13	0.25	0.19	Kumaraperumal et al., 2022	Eastern Districts of Tamil Nadu, India	RF
	1.27–1.47, 1.21, 1.36?		Liu et al., 2022b	China	Quantile regression forest
0.20–0.27, 0.15, 0.09	0.18, 0.23, 0.28		Mulder et al., 2016	France	Cubist
	4.7–10.6?		Nussbaum et al., 2018	Switzerland	RF, SOM not SOC
0.27–0.39, 0.19, 0.17	3.51–3.63, 3.45, 2.81?		Padarian et al., 2017	Chile	CART; bootstrapping
	3.97?		Poggio et al., 2021	Global	SoilGrids 2.0; Quantile regression forest
0.41	3.63?		Ramcharan et al., 2017	Conterminous USA	RF, gradient boosting
0.11–0.44	0.20–0.33, 0.17, 0.14	0.56–0.59; 0.42; 0.32	Reddy et al., 2021	India	Modified 3D-RK, RF
0.59	0.68		Suleymanov et al. (2023)	Saline steppe, Russia	RF; 0–20 cm
Phosphorus					
R²	RMSE (mg kg⁻¹)	CCC	Source	Study Area	
0.32, 0.29, 0.19	3.59, 3.38, 3.19		<i>This study</i>	<i>Namibia</i>	
0.12	4.3		Hengl et al., 2017b	Africa	Extractable by Mehlich 3 method; Ensemble RF & gradient boosting
0.49		0.654	Hengl et al., 2021	Africa	Mehlich3; Ensemble ML models*; ln-transformed
Total Nitrogen					
R²	RMSE (mg kg⁻¹)	CCC	Source	Study Area	
0.56, 0.76, 0.28	210, 130, 104		<i>This study</i>	<i>Namibia</i>	
	691		Hengl et al., 2015	Africa	AfSoilGrids250m; RF
0.66	558**		Hengl et al., 2017b	Africa	**Total organic N extractable by wet oxidation; Ensemble RF & gradient boosting
0.73		0.845	Hengl et al., 2021	Africa	Ensemble ML models*; ln-transformed
	914–1057, 813, 771		Liu et al., 2022b	China	Quantile regression forest
0.39	2700		Ramcharan et al., 2017	Conterminous USA	RF, gradient boosting
	250		Poggio et al., 2021	Global	SoilGrids 2.0; Quantile regression forest
Sand					
R²	RMSE (%)	CCC	Source	Study Area	
0.45, 0.37, 0.33	11.7, 13.4, 13.9		<i>This study</i>	<i>Namibia</i>	
0.235	23.0		Hengl et al., 2014	Global	SoilGrids1km; multiple linear regression
	15.9		Hengl et al., 2015	Africa	AfSoilGrids250m; RF

0.786	13.1		Hengl et al., 2017a	Global	SoilGrids250m 2016 update; Ensemble: RF, gradient boosting and/or multinomial logistic regression
0.736	13.7	0.848	Hengl et al., 2021	Africa	Ensemble ML models*
0.45–0.52, 0.48, 0.49	11.95–13.54, 13.29, 13.28		Liu et al., 2022a	Qinghai–Tibet Plateau, China	RF; 10-fold CV
0.53, 0.47, 0.47			Mgohele et al., 2024	[Review article]	RF; median values from ~31 studies
0.08–0.09, 0.05, 0.05	20.1–20.4, 21.9, 23.4		Padarian et al., 2017	Chile	CART; bootstrapping
0.57	17.8		Ramcharan et al., 2017	Conterminous USA	RF, gradient boosting
0.16–0.24	11.3–12.2, 11.4, 11.6	0.43–0.44; 0.40; 0.37	Reddy et al., 2021	India	Modified 3D-RK, RF
Silt					
R²	RMSE (%)	CCC	Source	Study Area	
0.43, 0.39, 0.36	7.0, 7.1, 7.3		<i>This study</i>	<i>Namibia</i>	
0.35	11.9		Hengl et al., 2014	Global	SoilGrids1km; multiple linear regression
	8.31		Hengl et al., 2015	Africa	AfSoilGrids250m; RF
0.79	9.8		Hengl et al., 2017a	Global	SoilGrids250m 2016 update; Ensemble: RF, gradient boosting and/or multinomial logistic regression
0.64	8.92	0.780	Hengl et al., 2021	Africa	Ensemble ML models*
0.47–0.50, 0.44, 0.49	8.64–9.11, 9.30, 8.77		Liu et al., 2022a	Qinghai–Tibet Plateau, China	RF; 10-fold CV
0.50, 0.46, 0.48			Mgohele et al., 2024	[Review article]	RF; median values from ~31 studies
	6.0–15.2		Nussbaum et al., 2018	Switzerland	RF
0.05–0.06, 0.03, 0.02	13.2–13.9, 14.1, 14.8		Padarian et al., 2017	Chile	CART; bootstrapping
Clay					
R²	RMSE (%)	CCC	Source	Study Area	
0.31, 0.23, 0.28	6.9, 9.1, 9.3		<i>This study</i>	<i>Namibia</i>	
0.24	18.1		Hengl et al., 2014	Global	SoilGrids1km; multiple linear regression
	13.7		Hengl et al., 2015	Africa	AfSoilGrids250m; RF
0.73	9.5		Hengl et al., 2017a	Global	SoilGrids250m 2016 update; Ensemble: RF, gradient boosting and/or multinomial logistic regression
0.75	9.6	0.854	Hengl et al., 2021	Africa	Ensemble ML models*
0.18–0.27	8.55–9.05	0.29–0.41	Kasraie et al., 2024	British Columbia, Canada	RF; 0–30 cm

0.30–0.36, 0.28,0.32	5.85–6.49, 6.91, 6.84		Liu et al., 2022a	Qinghai–Tibet Plateau, China	RF; 10-fold CV
0.49, 0.45, 0.43			Mgohele et al., 2024	[Review article]	RF; median values from ~31 studies
	6.1–8.9		Nussbaum et al., 2018	Switzerland	RF
0.04–0.06, 0.03, 0.03	12.5–14.1, 16.1, 17.1		Padarian et al., 2017	Chile	CART; bootstrapping
0.46	12.0		Ramcharan et al., 2017	Conterminous USA	RF, gradient boosting
0.35–0.40	6.9–7.2, 7.3, 6.7	0.52–0.55; 0.56; 0.50	Reddy et al., 2021	India	Modified 3D-RK, RF

Values separated by commas refer to different depth intervals; the same as this study (0–30, 30–60, and 60–100 cm), unless otherwise specified in last column.

Question marks mean that there was some uncertainty about the measuring units or conversion factors.

**Ensemble methods of Hengl et al. (2021): Random Forest, Extreme Gradient Boosting, Deep Learning, Cubist Regression Trees, and GLM with Lasso or Elasticnet Regularization*

U. NamSIS datasets and scripts

Soils of Namibia website

<https://soilsofnamibia.com/index.html>

This is a central access point to all NamSIS components. Development is still in progress.

Namibian Soil Profile Database

Coetzee, M., (2025). *The Namibian soil profile database (NSPD2025): an updated, expanded and harmonised compilation of national soil data (version 1)* [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.17618737>

DSM products

Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Base Saturation (%) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm) (1.0.0)* [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18832944>

Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Bulk Density (kg dm⁻³) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm) (1.0.0)* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.18832925>

Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Cation Exchange Capacity (cmolc kg⁻¹) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm) (1.0.0)* [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18833128>

Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Clay (%) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm) (1.0.0)* [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18833199>

Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Electrical Conductivity (1:2.5) (μS cm⁻¹) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm) (1.0.0)* [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18833300>

Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Electrical Conductivity (sat) (μS cm⁻¹) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm) (1.0.0)* [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18833395>

Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Extractable Calcium (mg kg⁻¹) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm) (1.0.0)* [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18833031>

Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Extractable Magnesium (mg kg⁻¹) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm) (1.0.0)* [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18833566>

Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Extractable Potassium (mg kg⁻¹) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm) (1.0.0)* [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18833488>

Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Extractable Sodium (mg kg⁻¹) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm) (1.0.0)* [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18833631>

- Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Organic Carbon (%) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm)* (1.0.0) [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18833807>
- Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted pH (H₂O) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm)* (1.0.0) [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18833833>
- Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Plant-available Phosphorus (mg kg⁻¹) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm)* (1.0.0) [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18833908>
- Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Sand (%) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm)* (1.0.0) [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18833968>
- Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Silt (%) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm)* (1.0.0) [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18834071>
- Coetzee, M., & Gelsleichter, Y. (2026). *NamSoil v1.0: Predicted Total Nitrogen (mg kg⁻¹) for Namibia at 90 m resolution (0–30, 30–60 and 60–100 cm)* (1.0.0) [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.18833711>

DSM scripts

- Gelsleichter, Y., & Coetzee, M. (2026). *NamSoil v1.0 – R and Google Earth Engine (GEE) code for digital soil mapping of Namibia at 90 m resolution* (1.0). [Software]. Zenodo. <https://doi.org/10.5281/zenodo.18776302>
- Gelsleichter, Y. (2026). Accessing NamSoil layers from Zenodo. [Software]. https://github.com/Gelsleichter/acquire_NamSoil/
- Gelsleichter, Y. (2026). Namibia Map (Version 1) [JavaScript]. Google Earth Engine. https://code.earthengine.google.com/?accept_repo=users/Namibia_map/Soil_properties

Namibian Soil Literature / Documents

Zotero Reference Management System

(https://www.zotero.org/groups/6649154/soil_literature_-_namibia/library)

Environmental Information Service (EIS) eLibrary (<http://the-eis.com/elibrary/>);

Namibia University of Science and Technology (NUST) Library Special Collections (<http://specialcollections.nust.na:8080/greenstone3/library/collection/soils/page/about>).

ACKNOWLEDGEMENTS

I express my sincere gratitude to my supervisors, Professors Erika Michéli and Adám Csorba of the Department of Soil Science, Doctoral School of Natural Sciences, Hungarian University of Agriculture and Life Sciences (MATE), for their guidance, encouragement, and support throughout this research. Their scientific insight and constructive criticism strengthened the study and helped me refine its conceptual and methodological foundations.

I thank the university colleagues and specialists, including the reviewers and MATE Department of Soil Science staff, who provided technical input, scientific advice, and practical assistance at different stages of the research. Their contributions strengthened the quality of the work and helped bring the dissertation to completion. I am particularly grateful to Dr Yuri Gelsleichter for assistance with the technical aspects of digital soil mapping. His expertise, practical advice, and assistance with data preparation, modelling procedures, programming, validation, and interpretation were invaluable to the development of the national soil property models, and his kindness and patience were boundless. My appreciation is also extended to the administrative staff of MATE, especially within the Department of Soil Science, the Doctoral and Habilitation Centre, and the International Student Relation Centre, who provided support in the many administrative requirements associated with the doctoral studies.

I gratefully acknowledge the Government of Hungary and the Tempus Public Foundation for awarding me the Stipendium Hungaricum Scholarship, which enabled me to undertake my doctoral studies. Their financial support provided the opportunity to complete this research and to contribute to the development of soil information for Namibia. I also highly appreciate the financial assistance with travel expenses provided by the Spatial Biodiversity Assessments, Prioritisation and Planning (SBAPP) project.

This research depended on the soil observations, maps, reports, photographs, laboratory records, and other legacy materials contributed by numerous organisations, projects, and individuals. Special acknowledgement is due to the Namibian Ministry of Agriculture, Fisheries, Water and Land Reform. The Ministry conducted the National Soil Survey almost three decades ago within their Agro-ecological Zoning Programme and granted permission for the resulting datasets to be used as the foundation of the present research. The foresight and work of those agricultural researchers, field and laboratory technicians created a national scientific resource whose value extends far beyond its small beginnings. I also thank the following organisations: Agri-Ecological Services, Gobabeb Namib Research Institute, and SASSCAL; projects: African Soil Microbial Genomics, Biodiversity Monitoring Transect Analysis in Africa (BIOTA), Assessment of the Impact of Bush Encroachment and Bush Control on Soil Organic Carbon in Namibia, Land Potential Knowledge System (LandPKS), EU Horizon 2020 Soils4Africa, The Future Okavango, Landscape Ecological Risk Assessment (LERIS), and Land Degradation Neutrality – Namibia; and individuals who shared their resources, including Josephath Kutuahupira; Dr Ben Strohhach, Dr Cornelis van der Waal, Dr John Mendelsohn, Dr Sophie Simmons, and Dr Jürgen Kempf.

I am thankful to my employer, the Namibian University of Science and Technology, for granting me study leave and my colleagues for their encouragement and support in making it possible to balance my studies with my professional commitments.

Finally, I offer my deepest thanks to my sons, Gian and Altus, family and friends for their support throughout this journey.