



Global Soil Nutrient and Nutrient Budgets maps (GSNmap) Phase I Technical Manual

Food and Agriculture Organization of the United Nations
Rome, 2022

Required citation:

FAO. 2022. *Global Soil Organic Carbon Map – GSOCmap v.1.6. Technical report*. Rome. <https://doi.org/10.4060/cb9015en>

The designations employed and the presentation of material in this information product do not imply the expression of any opinion whatsoever on the part of the Food and Agriculture Organization of the United Nations (FAO) concerning the legal or development status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. The mention of specific companies or products of manufacturers, whether or not these have been patented, does not imply that these have been endorsed or recommended by FAO in preference to others of a similar nature that are not mentioned.

The views expressed in this information product are those of the author(s) and do not necessarily reflect the views or policies of FAO.

ISBN 978-92-5-135899-3

© FAO, 2022



Some rights reserved. This work is made available under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 IGO licence (CC BY-NC-SA 3.0 IGO; <https://creativecommons.org/licenses/by-nc-sa/3.0/igo/legalcode>).

Under the terms of this licence, this work may be copied, redistributed and adapted for non-commercial purposes, provided that the work is appropriately cited. In any use of this work, there should be no suggestion that FAO endorses any specific organization, products or services. The use of the FAO logo is not permitted. If the work is adapted, then it must be licensed under the same or equivalent Creative Commons licence. If a translation of this work is created, it must include the following disclaimer along with the required citation: "This translation was not created by the Food and Agriculture Organization of the United Nations (FAO). FAO is not responsible for the content or accuracy of this translation. The original [Language] edition shall be the authoritative edition."

Disputes arising under the licence that cannot be settled amicably will be resolved by mediation and arbitration as described in Article 8 of the licence except as otherwise provided herein. The applicable mediation rules will be the mediation rules of the World Intellectual Property Organization <http://www.wipo.int/amc/en/mediation/rules> and any arbitration will be conducted in accordance with the Arbitration Rules of the United Nations Commission on International Trade Law (UNCITRAL).

Third-party materials. Users wishing to reuse material from this work that is attributed to a third party, such as tables, figures or images, are responsible for determining whether permission is needed for that reuse and for obtaining permission from the copyright holder. The risk of claims resulting from infringement of any third-party-owned component in the work rests solely with the user.

Sales, rights and licensing. FAO information products are available on the FAO website (www.fao.org/publications) and can be purchased through publications-sales@fao.org. Requests for commercial use should be submitted via: www.fao.org/contact-us/licence-request. Queries regarding rights and licensing should be submitted to: copyright@fao.org.

Cover design: ©FAO/Matteo Sala

Contents

Licence	ix
Abbreviations and acronyms	x
Contributors and reviewers	xi
Presentation and basics	xii
0.1 How to use this book	xii
0.2 Training material	xii
0.3 Setting up the software environment	xii
0.3.1 Digital Soil Mapping	xii
0.3.2 Main concepts: SCORPAN model	xii
Sampling design for soil surveys	xiii
0.4 Selection of the sampling methodology	xiii
0.5 Inference Methods	xiii
0.6 Purpose of Sampling	xiii
0.7 Sampling Design Types	xiii
0.7.1 Example: Soils4Africa Sampling Design	xiii

0.8	Soil Properties	xiii
0.9	Environmental Covariates	xiii
0.10	Understanding the Methodology	xiii
0.10.1	PSU Selection Process	xiv
0.10.2	SSU and TSU Selection Process	xiv
0.10.3	Site Identification System	xiv
0.11	Setting Up the Environment	xiv
0.11.1	Install Required Libraries	xiv
0.12	Define Variables and Parameters	xiv
0.12.1	Country ISO Code	xiv
0.12.2	Land Use Type	xiv
0.12.3	File Paths	xiv
0.12.4	Coordinate Reference System	xiv
0.12.5	Sample Size Calculation	xiv
0.12.6	Sampling Unit Definitions	xiv
0.12.7	Algorithm Parameters	xiv
0.13	Custom Functions	xiv
0.13.1	Covariate Space Coverage Function (CSIS)	xiv
0.13.2	TSU Generation Function	xiv
0.14	Load Country Boundaries and Legacy Data	xiv
0.14.1	Visualize Boundaries and Legacy Data	xv
0.15	Load Environmental Covariates for PSUs	xv
0.15.1	Optional Auxiliary Data	xv
0.15.2	Load and Process Auxiliary Data	xv
0.15.3	Load Main Environmental Covariates	xv
0.15.4	Load and Process Soil Climate Data (Newhall)	xv

0.15.5	Convert Categorical Variables to Dummy Variables	xv
0.15.6	Merge All Covariate Layers	xv
0.15.7	Crop and Save Covariates	xv
0.16	Principal Component Analysis (PCA)	xv
0.17	Organize the Sampling Universe	xv
0.17.1	Load and Process Land Use Data	xv
0.17.2	Exclude Protected Areas	xv
0.17.3	Restrict to Accessible Slopes	xv
0.17.4	Aggregate to 100m Resolution	xv
0.17.5	Save Processed Land Use	xv
0.17.6	Filter Legacy Data	xv
0.18	Create Primary Sampling Units (PSU Grid)	xvi
0.18.1	Generate 2km × 2km Grid	xvi
0.18.2	Trim to Country Boundaries	xvi
0.18.3	Save PSU Grid	xvi
0.19	Select PSUs with Sufficient Land Use Coverage	xvi
0.19.1	Calculate Land Use Percentage	xvi
0.19.2	Filter PSUs	xvi
0.20	Rasterize PSUs for Covariate Space Coverage	xvi
0.21	Compute Optimal Sample Size	xvi
0.21.1	Run Optimization	xvi
0.21.2	Extract Optimal Sample Size	xvi
0.22	Covariate Space Coverage - Computing PSUs	xvi
0.22.1	Prepare Function Parameters	xvi
0.22.2	Perform K-means Clustering	xvi
0.22.3	Assign Clusters and Calculate Distances	xvi

0.22.4	Extract Selected PSUs	xvi
0.22.5	Visualize in Covariate Space	xvii
0.23	Compute SSUs and TSUs	xvii
0.23.1	Load High-Resolution Covariates	xvii
0.23.2	Process Each PSU	xvii
0.23.3	Combine SSUs and TSUs	xvii
0.24	Export Sampling Units	xvii
0.24.1	Create Cluster Raster	xvii
0.24.2	Join Cluster Info to TSUs	xvii
0.24.3	Create Site IDs	xvii
0.24.4	Export Shapefiles	xvii
0.25	Compute Alternative PSUs	xvii
0.26	Assign Country-Wide Unique IDs	xvii
0.27	Export Unified Datasets	xvii
0.28	Summary Statistics	xvii
0.29	Create Distribution Maps	xvii
0.30	GPS Device Preparation	xviii
0.30.1	Export to GPX Format	xviii
0.30.2	Load onto GPS Device	xviii
0.31	Sampling Procedure	xviii
0.32	Data Recording	xviii
0.33	Quality Control	xviii
Appendix A: Software and Data Sources		xviii
0.33.1	Software Requirements	xviii
0.33.2	Data Sources	xviii
Appendix B: Troubleshooting Common Issues		xviii

0.33.3	Memory Issues	xviii
0.33.4	CRS Misalignment	xviii
0.33.5	Empty Geometry	xviii
	Appendix C: Acronyms and Abbreviations	xviii
TBD		xix
Soil Data Preparation		xx
0.34	Introduction	xx
0.35	Data cleaning	xx
0.36	Soil Data Harmonization	xx
0.37	Concepts on digital exchange of soil-related data (ISO 28258) . .	xx
0.38	Plot data	xx
0.39	Profile data	xx
0.40	Element data	xx
0.41	Specimen data	xx
0.42	Definition of laboratory procedures	xx
0.43	Data transformation	xxi
0.44	Organizing soil data according to the GloSIS database	xxi
0.45	Tools for harmonization of soil databases within GloSIS	xxi
Digital soil mapping and modeling		xxii
Soil Data Sharing		xxiii
0.46	Data sharing and export formats	xxiii
0.47	Metadata	xxiii
0.48	Web Services	xxiii
References		xxiv

Figures

Tables



Licence

The Technical Manual is made available under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 IGO licence

CC BY-NC-SA 3.0 IGO.

Abbreviations and acronyms

BD Bulk density

CEC Cation exchange capacity

CRAN Comprehensive R archive network

DSM Digital soil mapping

GEE Google Earth Engine

GSP Global Soil Partnership

INSII International Network of Soil Information Institutions

ITPS Intergovernmental Technical Panel on Soils

ME Mean error

MAE Mean absolute error

MEC Modelling efficiency coefficient

NDVI Normalized difference in vegetation index

QA/QC Quality assurance/quality check

RMSE Root mean squared error

SOC Soil organic carbon

SOM Soil organic matter

Contributors and reviewers

Presentation and basics

Placeholder

0.1 How to use this book

0.2 Training material

0.3 Setting up the software environment

0.3.1 Digital Soil Mapping

0.3.2 Main concepts: SCORPAN model

Sampling design for soil surveys

Placeholder

0.4 Selection of the sampling methodology

0.5 Inference Methods

0.6 Purpose of Sampling

0.7 Sampling Design Types

0.7.1 Example: Soils4Africa Sampling Design

0.8 Soil Properties

0.9 Environmental Covariates

0.10 Understanding the Methodology

- 0.10.1 PSU Selection Process
- 0.10.2 SSU and TSU Selection Process
- 0.10.3 Site Identification System
- 0.11 Setting Up the Environment**
 - 0.11.1 Install Required Libraries
- 0.12 Define Variables and Parameters**
 - 0.12.1 Country ISO Code
 - 0.12.2 Land Use Type
 - 0.12.3 File Paths
 - 0.12.4 Coordinate Reference System
 - 0.12.5 Sample Size Calculation
 - 0.12.6 Sampling Unit Definitions
 - 0.12.7 Algorithm Parameters
- 0.13 Custom Functions**
 - 0.13.1 Covariate Space Coverage Function (CSIS)
 - 0.13.2 TSU Generation Function
- 0.14 Load Country Boundaries and Legacy Data**

- 0.14.1 Visualize Boundaries and Legacy Data
- 0.15 Load Environmental Covariates for PSUs
 - 0.15.1 Optional Auxiliary Data
 - 0.15.2 Load and Process Auxiliary Data
 - 0.15.3 Load Main Environmental Covariates
 - 0.15.4 Load and Process Soil Climate Data (Newhall)
 - 0.15.5 Convert Categorical Variables to Dummy Variables
 - 0.15.6 Merge All Covariate Layers
 - 0.15.7 Crop and Save Covariates
- 0.16 Principal Component Analysis (PCA)
- 0.17 Organize the Sampling Universe
 - 0.17.1 Load and Process Land Use Data
 - 0.17.2 Exclude Protected Areas
 - 0.17.3 Restrict to Accessible Slopes
 - 0.17.4 Aggregate to 100m Resolution
 - 0.17.5 Save Processed Land Use
 - 0.17.6 Filter Legacy Data

- 0.18 Create Primary Sampling Units (PSU Grid)**
 - 0.18.1 Generate 2km × 2km Grid**
 - 0.18.2 Trim to Country Boundaries**
 - 0.18.3 Save PSU Grid**
- 0.19 Select PSUs with Sufficient Land Use Coverage**
 - 0.19.1 Calculate Land Use Percentage**
 - 0.19.2 Filter PSUs**
- 0.20 Rasterize PSUs for Covariate Space Coverage**
- 0.21 Compute Optimal Sample Size**
 - 0.21.1 Run Optimization**
 - 0.21.2 Extract Optimal Sample Size**
- 0.22 Covariate Space Coverage - Computing PSUs**
 - 0.22.1 Prepare Function Parameters**
 - 0.22.2 Perform K-means Clustering**
 - 0.22.3 Assign Clusters and Calculate Distances**
 - 0.22.4 Extract Selected PSUs**

- 0.22.5 Visualize in Covariate Space
- 0.23 Compute SSUs and TSUs
 - 0.23.1 Load High-Resolution Covariates
 - 0.23.2 Process Each PSU
 - 0.23.3 Combine SSUs and TSUs
- 0.24 Export Sampling Units
 - 0.24.1 Create Cluster Raster
 - 0.24.2 Join Cluster Info to TSUs
 - 0.24.3 Create Site IDs
 - 0.24.4 Export Shapefiles
- 0.25 Compute Alternative PSUs
- 0.26 Assign Country-Wide Unique IDs
- 0.27 Export Unified Datasets
- 0.28 Summary Statistics
- 0.29 Create Distribution Maps

0.30 GPS Device Preparation

0.30.1 Export to GPX Format

0.30.2 Load onto GPS Device

0.31 Sampling Procedure

0.32 Data Recording

0.33 Quality Control

Appendix A: Software and Data Sources

0.33.1 Software Requirements

0.33.2 Data Sources

Appendix B: Troubleshooting Common Issues

0.33.3 Memory Issues

0.33.4 CRS Misalignment

0.33.5 Empty Geometry

Appendix C: Acronyms and Abbreviations

TBD

Soil Data Preparation

0.34 Introduction

0.35 Data cleaning

0.36 Soil Data Harmonization

0.37 Concepts on digital exchange of soil-related data (ISO 28258)

0.38 Plot data

0.39 Profile data

0.40 Element data

0.41 Specimen data

0.42 Definition of laboratory procedures

0.43 Data transformation

0.44 Organizing soil data according to the GloSIS database

0.45 Tools for harmonization of soil databases within GloSIS

Digital soil mapping and modeling

Soil Data Sharing

0.46 Data sharing and export formats

0.47 Metadata

0.48 Web Services

References



The Global Soil Partnership (GSP) is a globally recognized mechanism established in 2012. Our mission is to position soils in the Global Agenda through collective action. Our key objectives are to promote Sustainable Soil Management (SSM) and improve soil governance to guarantee healthy and productive soils, and support the provision of essential ecosystem services towards food security and improved nutrition, climate change adaptation and mitigation, and sustainable development.



Australian Government
**Department of Agriculture,
Water and the Environment**



**Ministry of Finance of the
Russian Federation**



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation



**Rural Development
Administration**

