State of the art in global soil resource mapping

D G Rossiter
Adjunct Professor

Section of Soil & Crop Sciences, SIPS, CALS, Cornell University
https://www.css.cornell.edu/faculty/dgr2/

16-February-2023

Outline

Outline

Why **global** mapping?

- Global **models** using soil information are increasingly used
  - example: Community Earth System Model (CESM) from NCAR[1]
- Global **assessments** of soil condition are increasingly required
  - example: soil C stocks
  - example: Status of the World's Soil Resources[2]
- These need **consistent** and **spatially-explicit** global data

[1] https://www.cesm.ucar.edu/models

[2] https://www.fao.org/documents/card/en/c/c6814873-efc3-41db-b7d3-2081a10ede50/

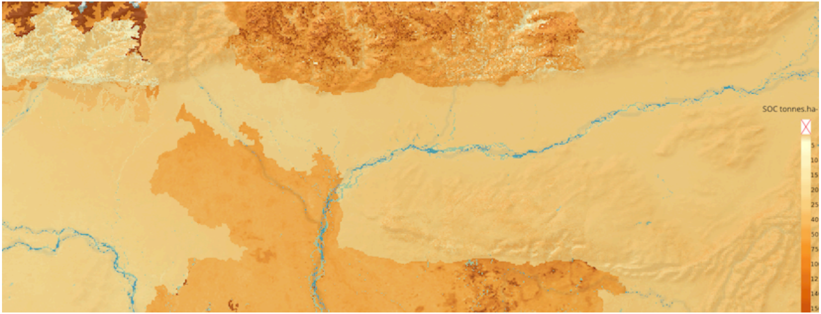
Why not combine national maps?

1. Many areas **not mapped** at any reasonable scale
2. Many other areas with **poor-quality** maps
3. Many areas **without key properties** needed by models, and/or without vertical distribution of the properties
4. **Inconsistency** in methods, definitions, mapping standards between adjacent countries

The FAO's attempt to stitch together national maps (1)

source: https://www.fao.org/global-soil-partnership/pillars-action/4-information-and-data-new/global-soil-organic-carbon-gsoc-map/en/

The FAO's attempt to stitch together national maps (2)



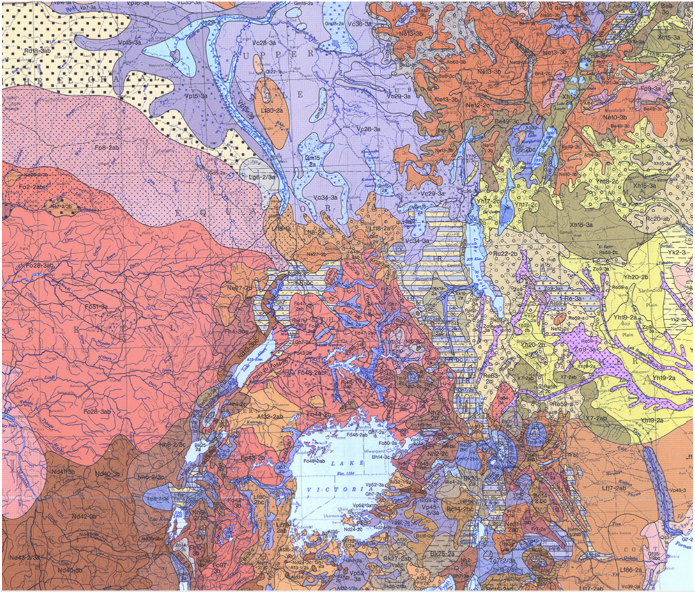
Outline

FAO Soil Map of the World

- 1961–1981 Soil Map of the World at 1:5 000 000[3]
  - soil **types** with a **unified legend** developed for the map, precursor to current international system World Reference Base (WRB)
  - broadly-defined soil classes based mainly on soil genesis, no systematic link to soil properties
  - the result of intensive "pedo-pedological" negotiations
  - many areas little more than **informed guesses**
  - rasterized by Zobler (1986), latest at $0.5°$ resolution; 106 soil units[4]

[3] https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/faounesco-soil-map-of-the-world/en/

[4] https://daac.ornl.gov/SOILS/guides/ZoblerSoilDerived.html


L A K E
V I C T O R I A
U P P E R
Fp8-2ab
Fo28-3ab
Fo28-3ab
Vc28-3a
Vc29-3a
Xh15-3a
Yh20-2b
Yh19-2a
X7-2ab
Lf17-2ab


SOIL UNITS *
UNITES PEDOLOGIQUES *
UNIDADES DE SUELO *
ПОЧВЕННЫЕ ЕДИНИЦЫ *
J FLUVISOLS FLUVISOLS FLUVISOLES ФЛЮВИСОЛИ
G GLEYSOLS GLEYSOLS GLEYSOLES ГЛЕЙСОЛИ
R REGOSOLS REGOSOLS REGOSOLES РЕГОСОЛИ
I LITHOSOLS LITHOSOLS LITOSOLES ЛИТОСОЛИ
Q ARENOSOLS ARENOSOLS ARENOSOLES АРЕНОСОЛИ
E RENDZINAS RENDZINES RENDZINAS РЕНДЗИНЫ
U RANKERS RANKERS RANKERS РАНКЕРЫ
T ANDOSOLS ANDOSOLS ANDOSOLES АНДОСОЛИ
V VERTISOLS VERTISOLS VERTISOLES ВЕРТИСОЛИ
Z SOLONCHAKS SOLONCHAKS SOLONCHAKS СОЛОНЧАКИ
S SOLONETZ SOLONETZ SOLONETZ СОЛОНЦЫ
Y YERMOSOLS YERMOSOLS YERMOSOLES ЕРМОСОЛИ
X XEROSOLS XEROSOLS XEROSOLES КСЕРОСОЛИ
K KASTANOZEMS KASTANOZEMS KASTANOZEMS КАШТАНОЗЕМЫ
H PHAEOZEMS PHAEOZEMS PHAEOZEMS ФАЙОЗЕМЫ
B CAMBISOLS CAMBISOLS CAMBISOLES КАМБИСОЛИ
L LUVISOLS LUVISOLS LUVISOLES ЛЮВИСОЛИ
P PODZOLS PODZOLS PODZOLES ПОДЗОЛЫ
W PLANOSOLS PLANOSOLS PLANOSOLES ПЛАНОСОЛИ
A ACRISOLS ACRISOLS ACRISOLES АКРИСОЛИ
N NITOSOLS NITOSOLS NITOSOLES НИТОСОЛИ
F FERRALSOLS FERRALSOLS FERRALSOLES ФЕРРАЛЬСОЛИ
O HISTOSOLS HISTOSOLS HISTOSOLES ГИСТОСОЛИ

Digitial Soil Map of the World

- FAO digitized the paper maps of its 1:5 000 000 map
- available as ESRI shapefile, Erdas, IDRISI GIS formats
- `https://data.apps.fao.org/map/catalog/srv/eng/catalog.search#/metadata/446ed430-8383-11db-b9b2-000d939bc5d8`

Harmonized World Soil Database v1.2 (HWSD)

- FAO, IIASA, ISRIC-World Soil Information, Institute of Soil Science, Chinese Academy of Sciences (ISSCAS), the Joint Research Centre of the European Commission (JRC)[5]
- 30' ($\approx$ 0.92 km horizontal resolution at Equator) raster database ($\approx\equiv 1{:}150\text{k}$ polygon map scale)
- composition within raster cells in terms of soil units (FAO revised legend)
- estimates of **soil property** values for **2 layers**
  - organic C, pH, water storage capacity, soil depth, cation exchange capacity of the soil and the clay fraction, total exchangeable nutrients, lime and gypsum contents, Na exchange percentage, salinity, textural class and particle-size separates)

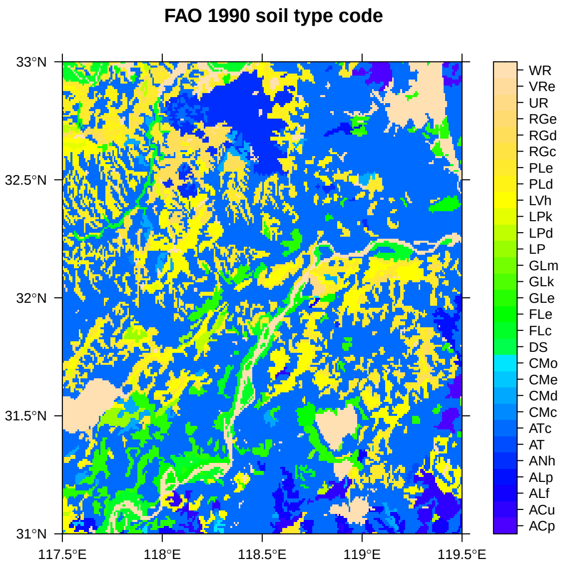
[5] https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/

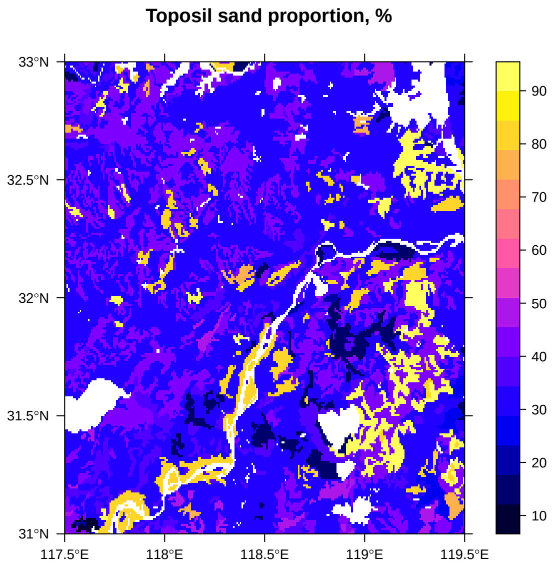
HWSD data fields

```
> dbGetQuery(con, "pragma table_info(HWSD_DATA)")$name

 [1] "ID"                "MU_GLOBAL"
 [3] "MU_SOURCE1"        "MU_SOURCE2"
 [5] "ISSOIL"            "SHARE"
 [7] "SEQ"               "SU_SYM74"
 [9] "SU_CODE74"         "SU_SYM85"
[11] "SU_CODE85"         "SU_SYM90"
[13] "SU_CODE90"         "T_TEXTURE"
[15] "DRAINAGE"          "REF_DEPTH"
[17] "AWC_CLASS"         "PHASE1"
[19] "PHASE2"            "ROOTS"
[21] "IL"                "SWR"
[23] "ADD_PROP"          "T_GRAVEL"
[25] "T_SAND"            "T_SILT"
[27] "T_CLAY"            "T_USDA_TEX_CLASS"
[29] "T_REF_BULK_DENSITY" "T_BULK_DENSITY"
[31] "T_OC"              "T_PH_H2O"
[33] "T_CEC_CLAY"        "T_CEC_SOIL"
[35] "T_BS"              "T_TEB"
[37] "T_CACO3"           "T_CASO4"
[39] "T_ESP"             "T_ECE"
[41] "S_GRAVEL"          "S_SAND"
[43] "S_SILT"            "S_CLAY"
[45] "S_USDA_TEX_CLASS"  "S_REF_BULK_DENSITY"
[47] "S_BULK_DENSITY"    "S_OC"
[49] "S_PH_H2O"          "S_CEC_CLAY"
[51] "S_CEC_SOIL"        "S_BS"
[53] "S_TEB"             "S_CACO3"
[55] "S_CASO4"           "S_ESP"
[57] "S_ECE"
```

MU = map unit
SU = soil unit
T = topsoil
S = subsoil


FAO 1990 soil type code
33°N
32.5°N
32°N
31.5°N
31°N
117.5°E
118°E
118.5°E
119°E
119.5°E
WR
VRe
UR
RGe
RGd
RGc
PLe
PLd
LVh
LPk
LPd
LP
GLm
GLk
GLe
FLe
FLc
DS
CMo
CMe
CMd
CMc
ATc
AT
ANh
ALp
ALf
ACu
ACp


Toposil sand proportion, %
33°N
32.5°N
32°N
31.5°N
31°N
117.5°E
118°E
118.5°E
119°E
119.5°E
90
80
70
60
50
40
30
20
10

HWSD v2.0

- Released late Jan-2023
  - FAO; IIASA. (2023). Harmonized World Soil Database version 2.0. FAO; International Institute for Applied Systems Analysis (IIASA); `https://doi.org/10.4060/cc3823en`
- Improvements in data sources, layer information (7 instead of 2), uniform soil classification (FAO 1990, WRB 2022) including Technosols.
- added properties effective CEC, total N, C/N ratio, Al saturation.
- "Its main objective is to be useful for modelers and to serve as a basis for prospective studies on agroecological zoning, food security and the impacts of climate change."

GlobalSoilMap.net

- A project from 2008, challenge to global soil mappers by Pedro Sánchez (CALS graduate, author of a standard work on soil fertility management of tropical soils)
- Gates Foundation financing
- Idea: **regional nodes** working to **one standard**
- Standards were successfully developed and widely-adopted by many mapping projects[6]
    - 90 m resolution grid; 7 depth slices; required soil properties; **uncertainty**
- **Project as such failed**, the regional nodes concept was unworkable outside of North America and Europe

[6] https://www.isric.org/sites/default/files/GlobalSoilMap_specifications_december_2015_2.pdf

GSM Standards

Specifications

Tiered[1] *GlobalSoilMap* products

Release 2.4

[07/12/2015][2]

Science Committee

[1] Tier 1 – point predictions at centre of a global3 arc-second grid; Tier 2 - additional 100mx100m block predictions centred on 3 arc-second grid cells

[2] These specifications have been modified as a result of an Uncertainty Workshop held at USDA NRCS, Lincoln, Nebraska, August 27-30, 2012 and an ad hoc meeting in Jeju, Korea, June 11, 2014.

1

GSM results: Australia

858 *Soil Research* R. A. Viscarra Rossel *et al.*

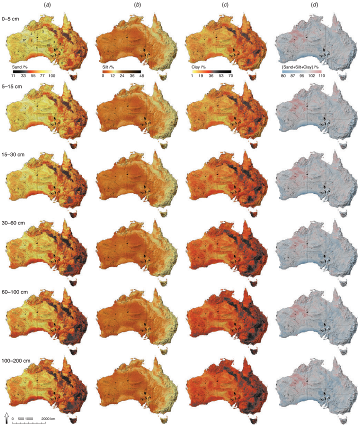
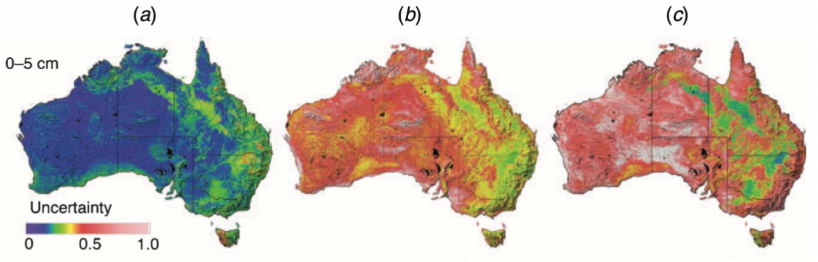


Fig. 5. Maps of (*a*) sand, (*b*) silt and (*c*) clay content for each of the six *GlobalSoilMap* standard depths. The sum of the sand, silt and clay contents for each standard depth appears in (*d*). In the map of sand content for the 0–5 cm, we shows the approximate location of Australian deserts: (1) Great Victoria, (2) Great Sandy, (3) Tanami, (4) Simpson, (5) Gibson, (6) Little Sandy, (7) Strzelecki, (8) Sturt Stony, (9) Tirari, (10) Pedirka, (11) Mallee.

Viscarra Rossel, R. A., *et al.*(2015). The Australian three-dimensional soil grid: Australia's contribution to the GlobalSoilMap project. Soil Research, 53(8), 845–864. `https://doi.org/10.1071/SR14366`

Uncertainty



Maps of the uncertainties in estimates of (a) sand, (b) silt and (c) clay content for each depth, expressed as the range of the 90% con!dence intervals standardised by their respective means.

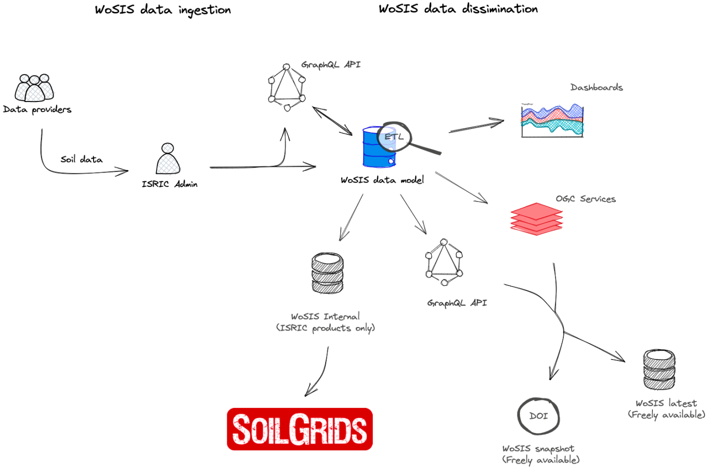
Outline

Data provider: ISRIC–World Soil Information

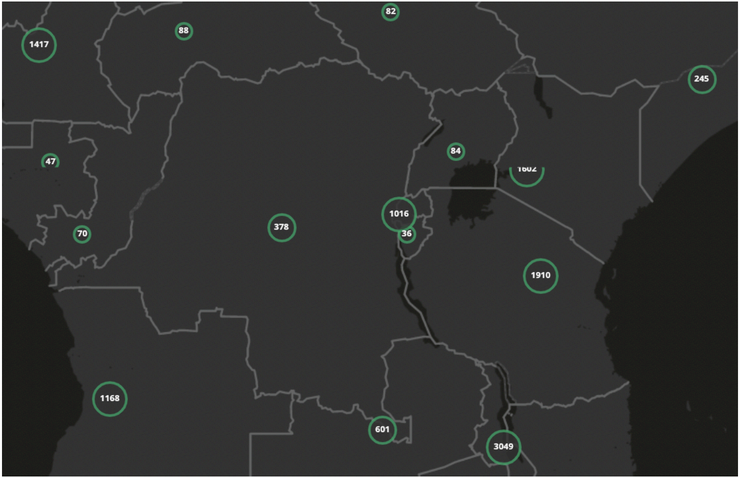
- "an independent foundation with a mission to serve the international community as a custodian of global soil information"
  - World Data Centre for Soils, part of the International Science Council (ISC) World Data System
- Home page `https://www.isric.org`
- harmonized point database (WoSIS)[7]
- global-extent soil maps (SoilGrids)[8]
- consistent point data + covariates → global models → **consistent global maps**

[7] `https://www.isric.org/explore/wosis`
[8] `https://www.isric.org/explore/soilgrids`

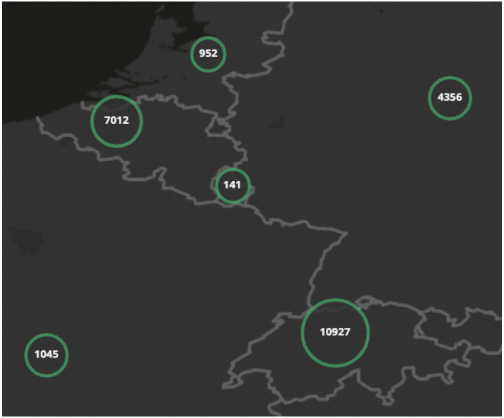
World Soil Information System (WoSIS)



Publically-available WoSIS profiles – Central Africa



Publically-available WoSIS profiles – NW Europe



SoilGrids v2.0

- GlobalSoilMap.net specifications, 250 m horizontal resolution
  - 7 depth slices; 5 physical properties; 4 chemical properties; 2 derived properties
  - with model uncertainty
  - most probable WRB Reference Soil Groups – outdated, in revision
  - "soon": more derived properties, soil functions
- also generalized products at 1 km, 5 km averages
- Online[9]: view, download
- Specifications, FAQ[10]
- Also available in Google Earth Engine catalog

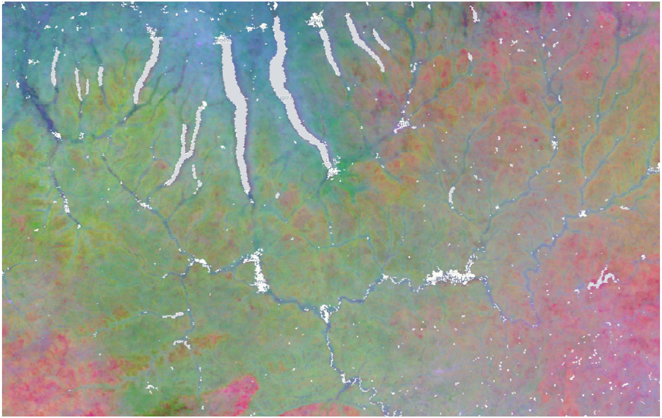
[9] https://www.soilgrids.org

[10] https://www.isric.org/explore/soilgrids/faq-soilgrids

Derived property: SOC stocks

SOC stocks model uncertainty [(Q0.95-Q0.05)/Q0.50] x 10

"Natural" clusters from SGv2.0 properties and layers



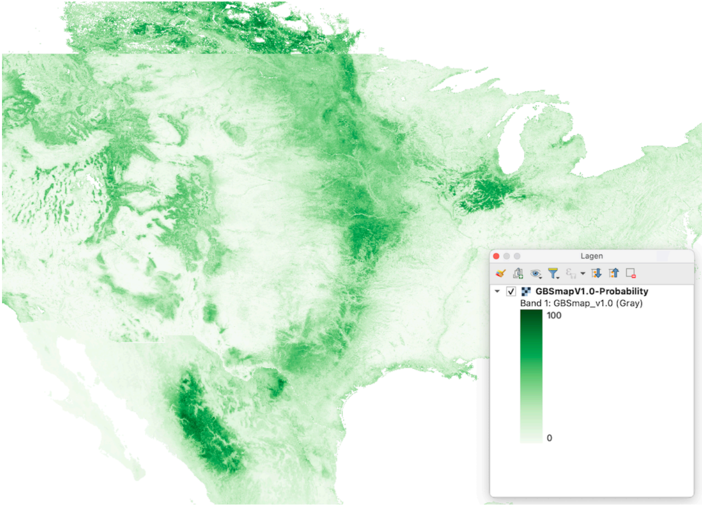
Google Earth Engine, Python API, R `reticulate` interface, R `rgee` package

Data provider: FAO Soil Resources

- Land & Water Division[11]; Soils office[12]
  - Global Soil Partnership[13] "improve the governance and promote sustainable management of soils", including mapping projects
  - Intergovernmental Technical Panel on Soils[14] advises GSP
- working with countries, stitching together the products
- recent products: SOC, "Black" soils[15]

[11] https://www.fao.org/land-water/home/en/
[12] https://www.fao.org/land-water/land/httpwwwfaoorgsoils-portalen/en/
[13] https://www.fao.org/global-soil-partnership
[14] https://www.fao.org/global-soil-partnership/intergovernmental-technical-panel-soils/en/
[15] https://www.fao.org/global-soil-partnership/gbsmap/en/

Probability of "Black" soil occurrence



Data provider: Sun Yat-sen University 中山大学

- Land-Atmosphere Interaction Research Group
- Global Soil Dataset for Earth System Modeling (GSDE)[16]
  - based on the Soil Map of the World and various regional and national soil database
  - soil type linkage method (i.e. taxotransfer rules) and the polygon linkage method to derive the spatial distribution of soil properties.
  - 8 layers, 30" grid cell resolution
- Global high-resolution (30", $\approx$ 1 km at equator) dataset of soil **hydraulic** and **thermal** parameters for land surface modeling[17]
  - derived respectively from the Global Soil Dataset for Earth System Models (GSDE) [Shangguan et al., 2014] and SoilGrids [Hengl et al., 2014, 2017] databases.
  - also generalized to 5' resolution

[16] http://globalchange.bnu.edu.cn/research/soilw

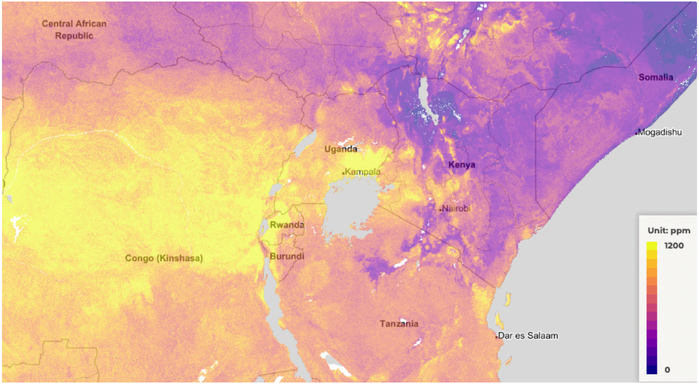
[17] http://globalchange.bnu.edu.cn/research/soil5.jsp

Data provider: EnvironmetriX

- Commercial company[18] based in Wageningen (NL), founded by Tomislav Hengl (SoilGrids v1)
- Because of demand, mostly regional products, e.g., iSDA 30 m resolution soil property maps (2 layers) for Africa[19]
- Linked to OpenGeoHub[20] non-profit foundation for geodata

[18] https://envirometrix.nl/
[19] https://www.isda-africa.com/isdasoil/
[20] https://opengeohub.org/

iSDA: extractable Fe

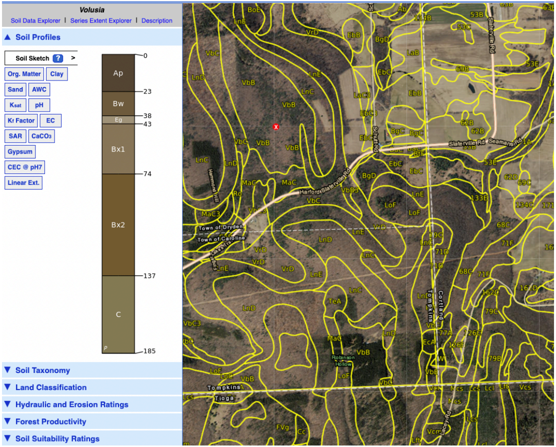


Outline

Global maps from legacy maps

- Examples: FAO Soil Map of the World, HSWD, gSSURGO (USA)
- Decide on scale/resolution, legend
  - restricted by source maps
- Compile diverse legacy maps to a **common legend**
- **Digitize** (if paper), **simplify/correlate/harmonize** (if already digital)
- Link to **database**: soil type, soil properties

gSSURGO shown in SoilWeb



source: https://casoilresource.lawr.ucdavis.edu/gmap/

Global maps from Digital Soil Mapping (DSM)

- Definition: making **direct-to-digital** soil maps at some grid resolution
  - McBratney, A. B., Mendonça Santos, M. L., & Minasny, B. (2003). On digital soil mapping. Geoderma, 117, 3–52. `https://doi.org/10.1016/S0016-7061(03)00223-4`
- **Source** information: **georeferenced pedon** observations: location + attributes
- **Output**: full-field gridded maps of soil properties or soil classes

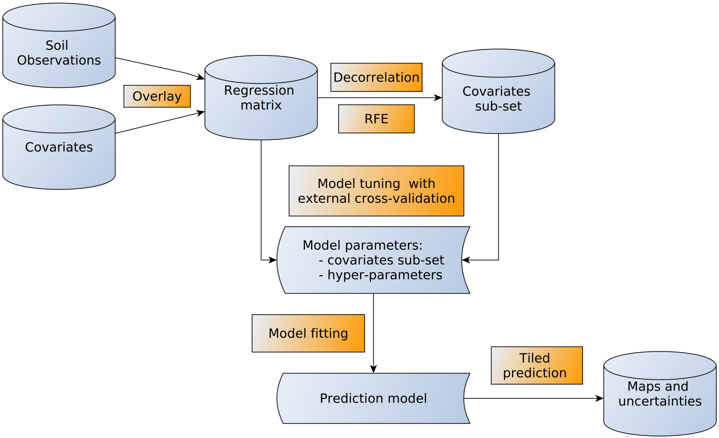
DSM methods

- **Geostatistical** methods (e.g., **kriging** from point observations) used in local applications (e.g., precision farming)
- **Machine learning** methods using georeferenced full-field **covariates** related to soil formation
  - covariates represent soil-forming factors
  - **empirical relation** between covariate values at points and the soil property values at the same points
  - most used method is **random forest**; also can use **gradient boosting**
  - model can then be computed at all grid cells → digital soil map

Why is DSM so popular?

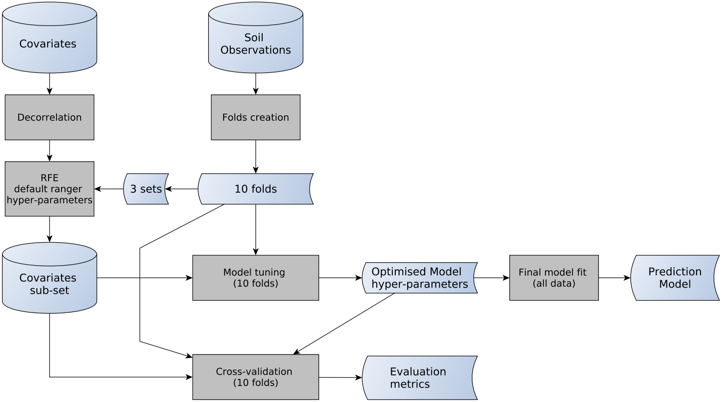
- relatively easy, rapid and cheap to produce a nice-looking product
- increasing **computer power**
  - Google Earth Engine; R, `ranger`, `xgboost`, `caret`...
- increasing availability of relevant **covariates**
  - DEM, satellite imagery, γ-rays, land surface temperature...
- good training materials, ready-made solutions (FAO, ISRIC)
- **little knowledge** of soil-landscape relations or soil genesis or is required
  - Lack of trained soil mappers, correlators, quality control procedures
  - Under-funded or non-existent soil survey organizations

SoilGrids 2.0 DSM workflow



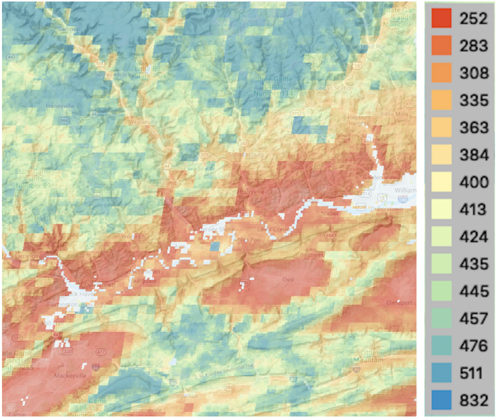
source: `https://dx.doi.org/10.5194/soil-7-217-2021`

SoilGrids 2.0 DSM model tuning



source: `https://dx.doi.org/10.5194/soil-7-217-2021`

Example SoilGrids 2.0 map fragment

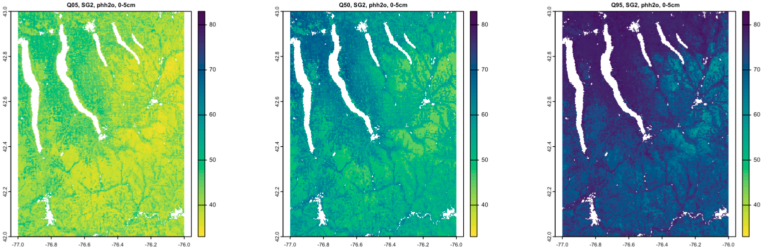


Sand, $\mathrm{g\ kg^{-1}}$, 0–5 cm layer, centres of 250 m resolution grid cells, near Jersey Shore, PA

Model uncertainty

- A big advantage of DSM: "known unknowns" (Donald Rumsfeld)
- Kriging: kriging prediction variance → confidence intervals
- Quantile random forests: can slice at any quantile (cumulative proportion of trees predicting at or below the quantile)
- Probability random forests: probability of each class
    - summarize as maximum probability, confusion index, Shannon entropy ...
- can refuse to map areas that are "too" uncertain for user requirements

SoilGrids v2.0 prediction quantiles Q05, Q50, Q95



pH x 10, 0–5 cm layer, centres of 250 m resolution grid cells

Outline

Portals

- My site of all Soil Geographic Databases, hosted by ISRIC-World Soil Information[21]
  - Global coverages, including other portals[22]
    - FAO, Global Soil Information System (GloSIS), World Data System, Global Earth Observation System of Systems, Distributed Active Archive Center (Oak Ridge) ...
  - Regional, national, local databases; point, grid, polygon formats

[21] https://www.isric.org/explore/soil-geographic-databases

[22] https://www.isric.org/explore/soil-geographic-databases#world

Outline

What are the main problems with these products?

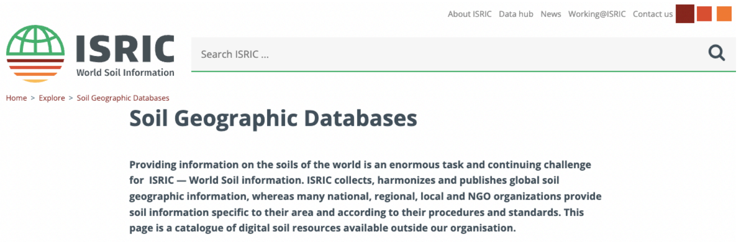
- Poor accuracy / **high uncertainty**
- **incomplete distribution** of points in the **feature space** used to make the predictive models (DSM)
- **incompatible / outdated / undocumented** calibration points, even in well-covered areas
- few maps of **derived properties** (e.g., soil hydrology, erodibility)
- few maps of **soil functions**

Future developments

- Larger, better-harmonized pedon databases
  - especially for areas poorly-represented in covariate space
- More covariates related to soil-forming processes
- Making maps of **derived properties** directly useful to modellers
- Making maps of **soil functions** directly useful in soil assessment
- More attention to communicating uncertainty and making maps as needed by the global community
  - e.g., identifying the **"area of applicability"**[23] and areas with predictions that are too uncertain for use

[23] Meyer, H., & Pebesma, E. (2020). Predicting into unknown space? Estimating the area of applicability of spatial prediction models. arXiv:2005.07939 `http://arxiv.org/abs/2005.07939`

End



https://www.isric.org/explore/soil-geographic-databases