

Preparing Environmental Covariates for Digital Soil Mapping

Ruhollah Taghizadeh
Soil Science and Geomorphology

E-mail: ruhollah.taghizadeh-mehrjardi@mnf.uni-tuebingen.de

Content

❑ Part 1

- Digital Soil Mapping
- Environmental Covariates

❑ Part 2

- Lets practice!

SCORPAN model

$$S = f(s, c, o, r, p, a, n) + \varepsilon$$

S : Soil, at a specific point in space and time: soil classes, **Sc** or soil attributes, **Sa**

From Jenny's Equation

c : climate, climate properties of the environment;

o : organisms, vegetation;

r : topography, landscape attributes;

p : parent material, lithology;

a : age or time factor;

Additions:

s : soil, prior knowledge of the soil at a point;

n : space, relative spatial position;

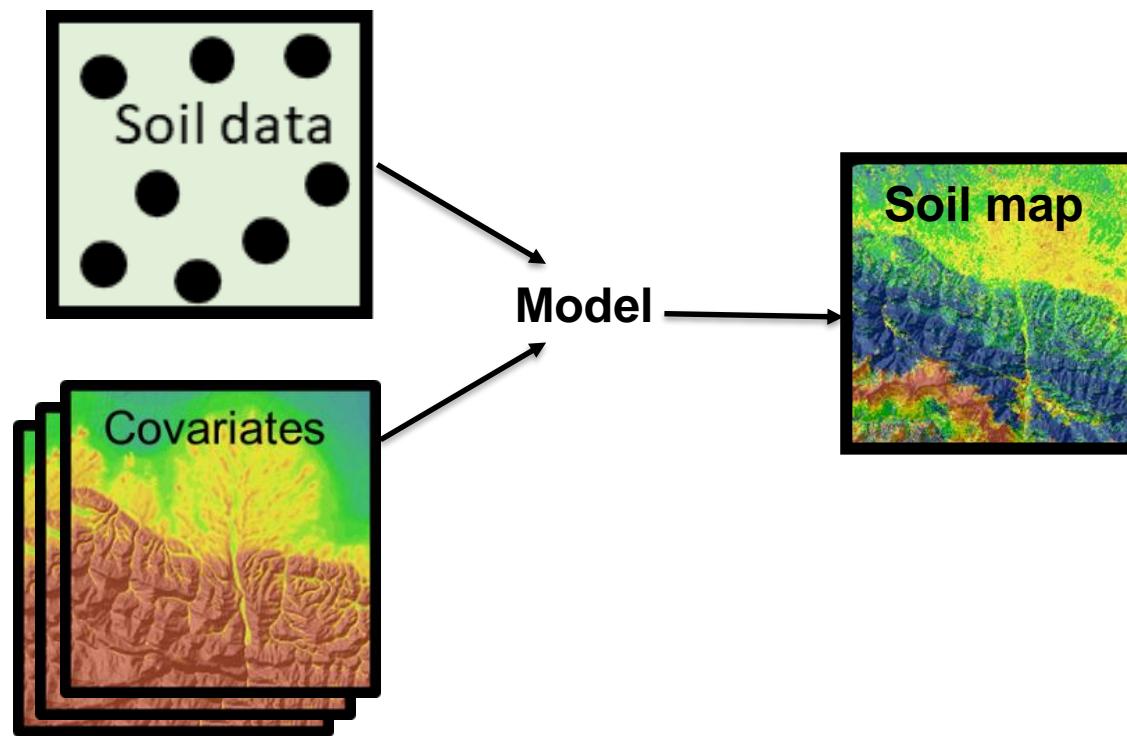
ε : auto-correlated random spatial variation.

f() : Quantitative function **f** linking **S** to **scorpan** factors

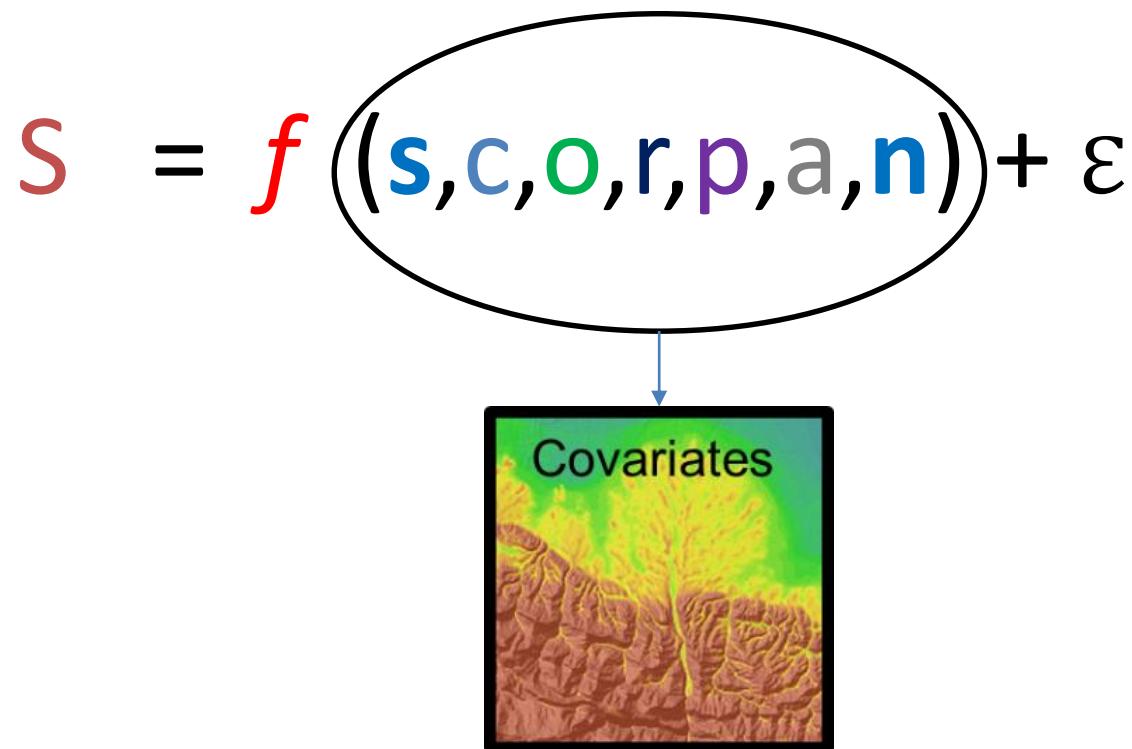
SCORPAN model

$$S = f(s, c, o, r, p, a, n) + \varepsilon$$

Soil data Model Covariates spatially dependent residuals



SCORPAN Model



Environmental Covariates

s

Possible sources of information to represent seven scorpan factors

c

scorpan factor Possible representatives

o

s Legacy soil maps, point observations, expert knowledge

r

c Temperature and precipitation records

p

o Vegetation maps, species abundance maps, yield maps, land use maps

a

r Digital elevation model, terrain attributes

n

p Legacy geology maps, gamma radiometric information

a Weathering indices, geology maps

n Latitude and longitude or easting and northing, distance from landscape features, distance from roads, distance from point sources of pollution

- Proximal sensing, Remote sensing, and digital elevation models are three prominent examples of high-resolution environmental covariates can be used to represent various scorpan factors

Proximal Soil Sensing (PSS)

s
c
o
r
p
a
n

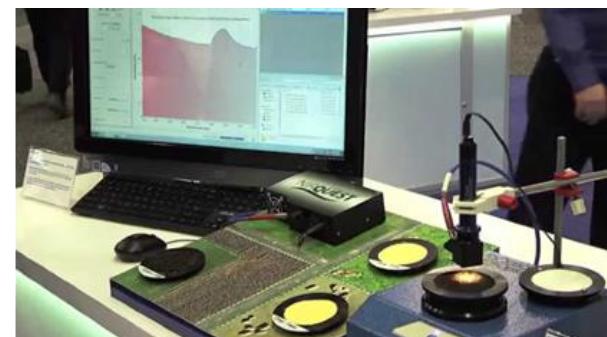
- Proximal soil sensing is a set of techniques developed to measure the physical, chemical and biological properties of soil when placing the sensor in contact, or at a proximal distance (within 2 m) to, the soil being characterized (Viscarra Rossel et al., 2011)
- The proximal soil sensors are more time- and cost effective than conventional laboratory analyses
- The proximal soil sensors facilitate the collection of larger amounts of (spatial) data



EM38



PXRF

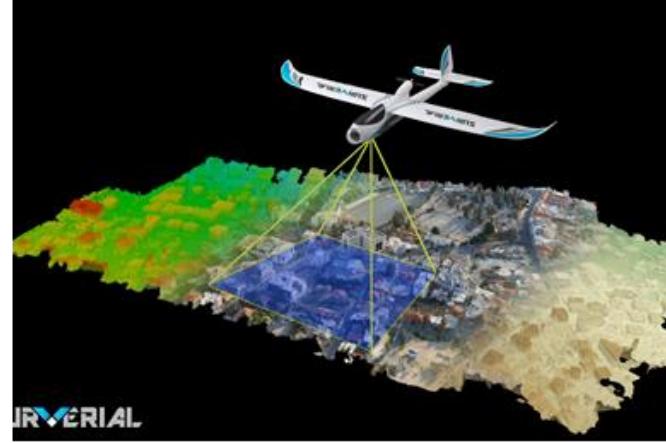
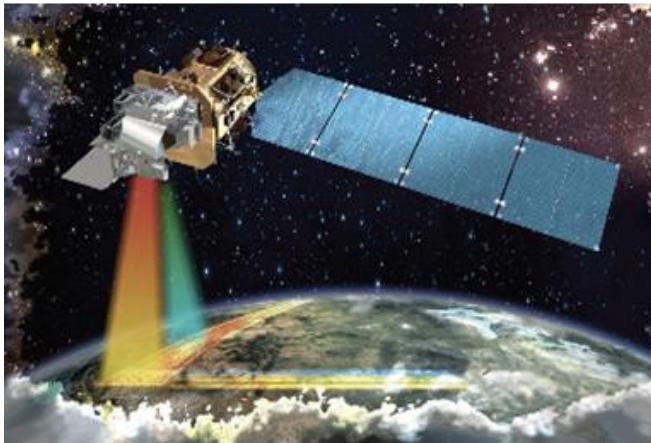


NIR

Remote Sensing

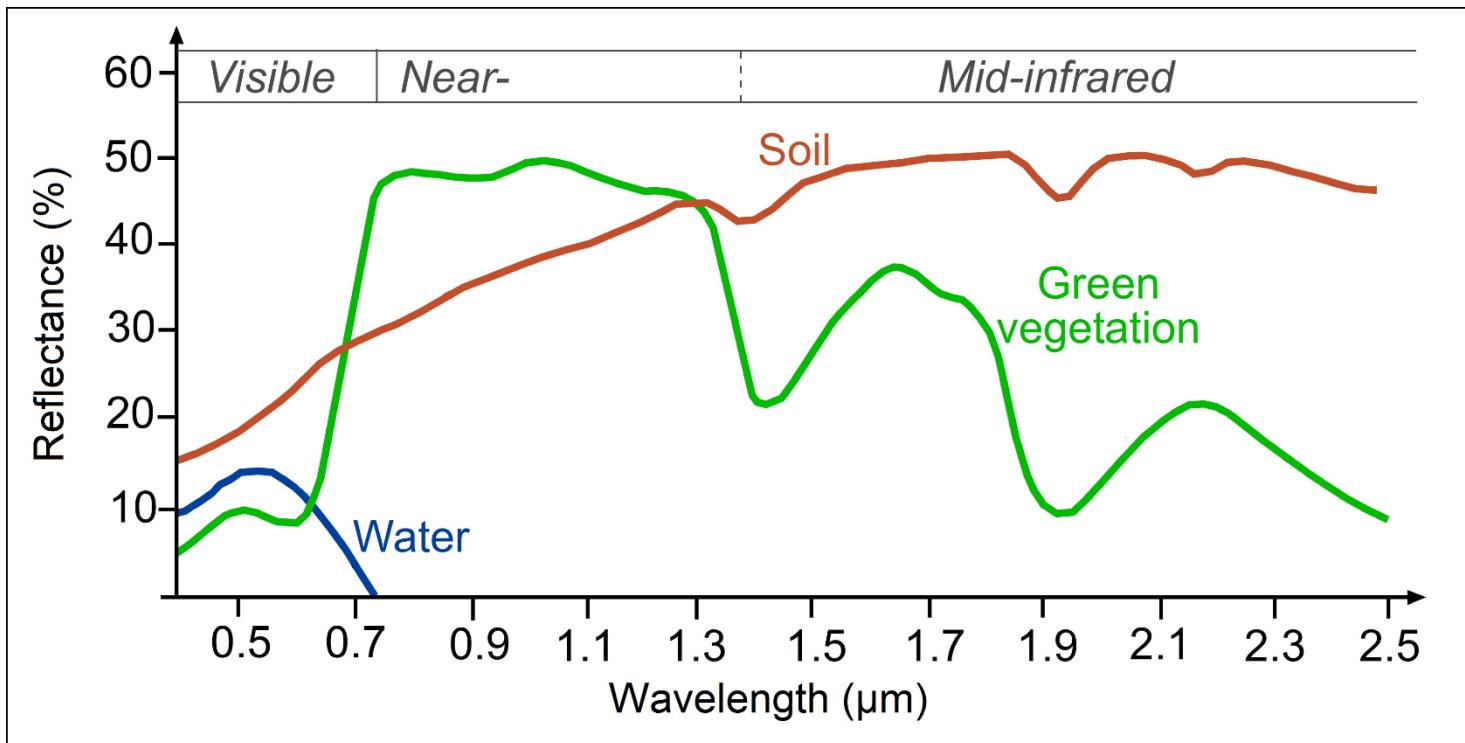
- The remotely sensed data has been shown to be an efficient means of assessing the condition of natural resources

s
c
o
r
p
a
n



Remote Sensing

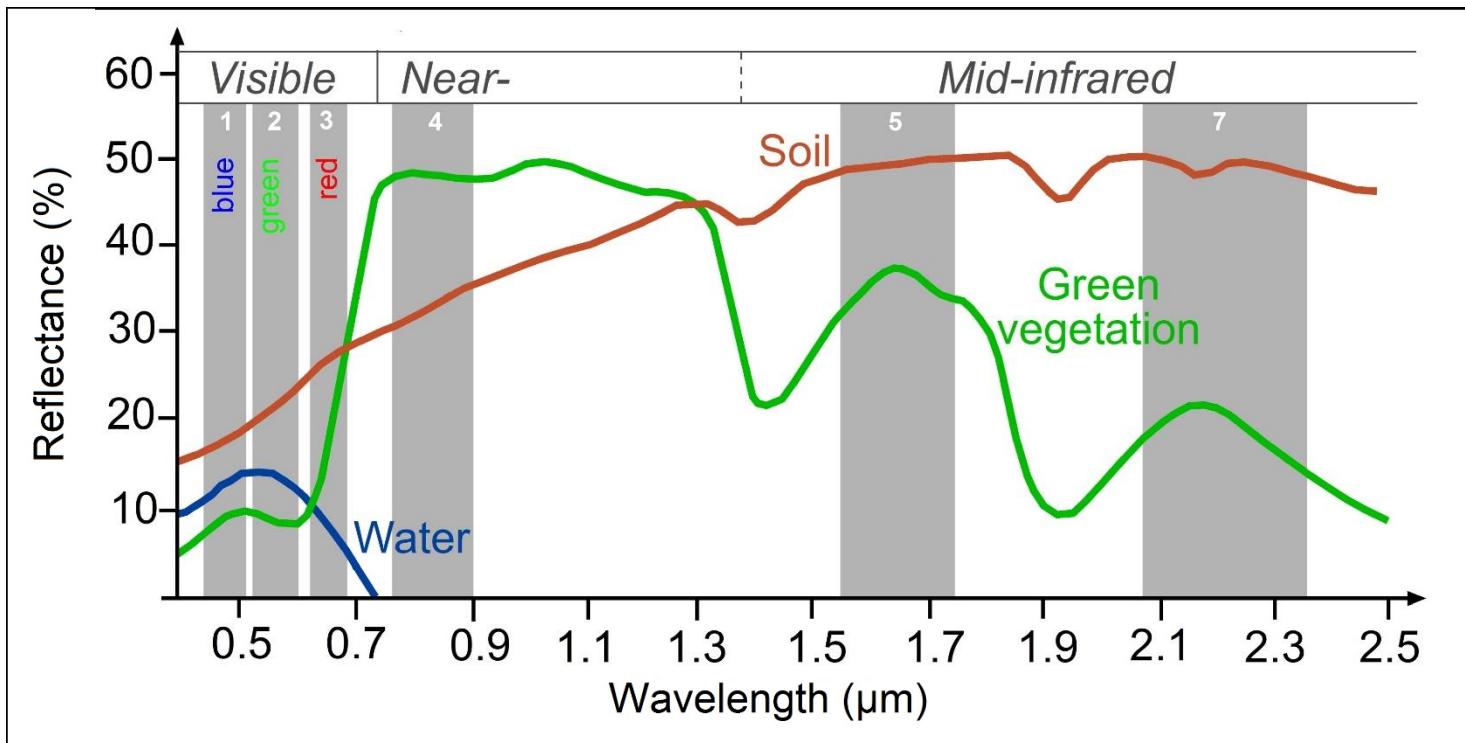
s
c
o
r
p
a
n



Spectral signatures of soil, vegetation and water.
Source: Siegmund, Menz 2005 with modifications

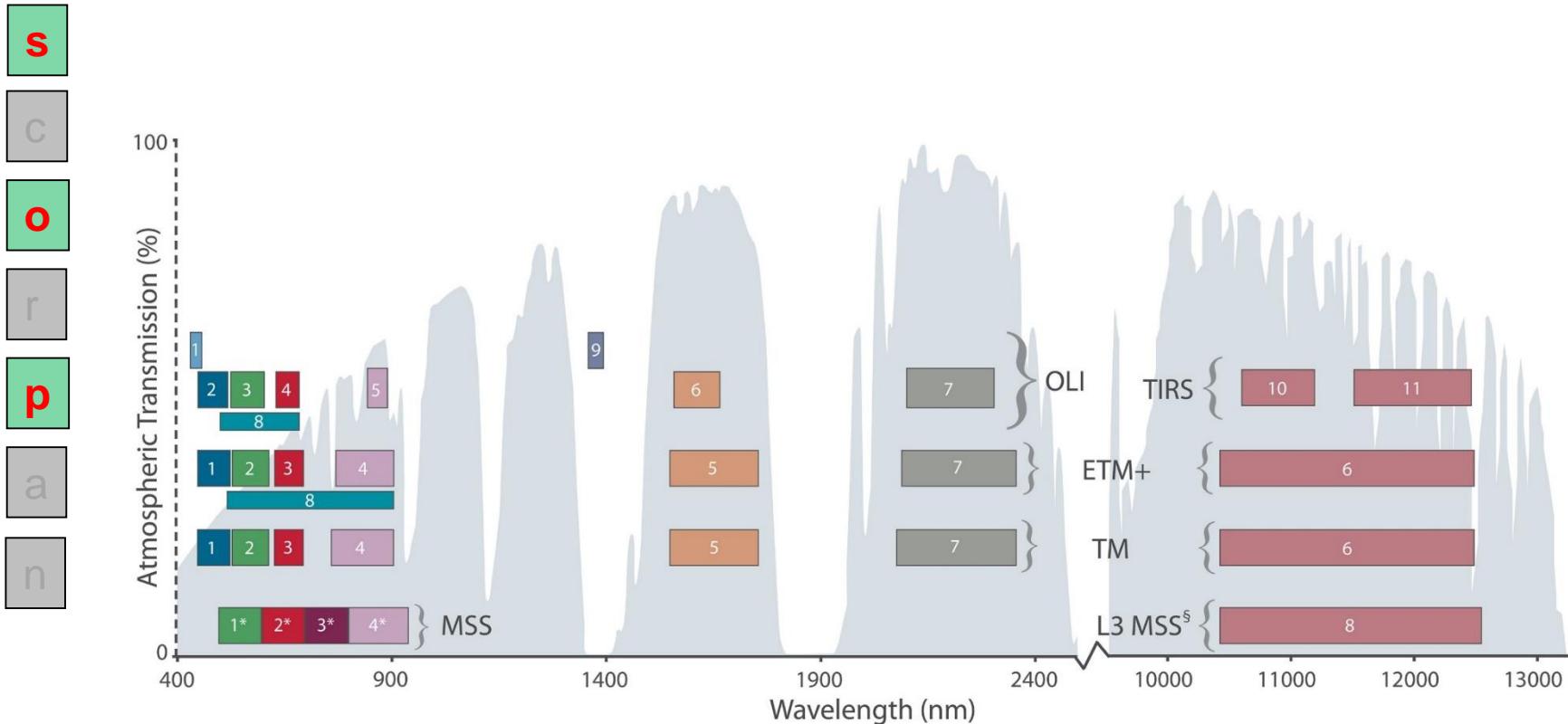
Remote Sensing

s
c
o
r
p
a
n



Spectral signatures of soil, vegetation and water, and spectral bands of LANDSAT 7.
Source: Siegmund, Menz 2005 with modifications

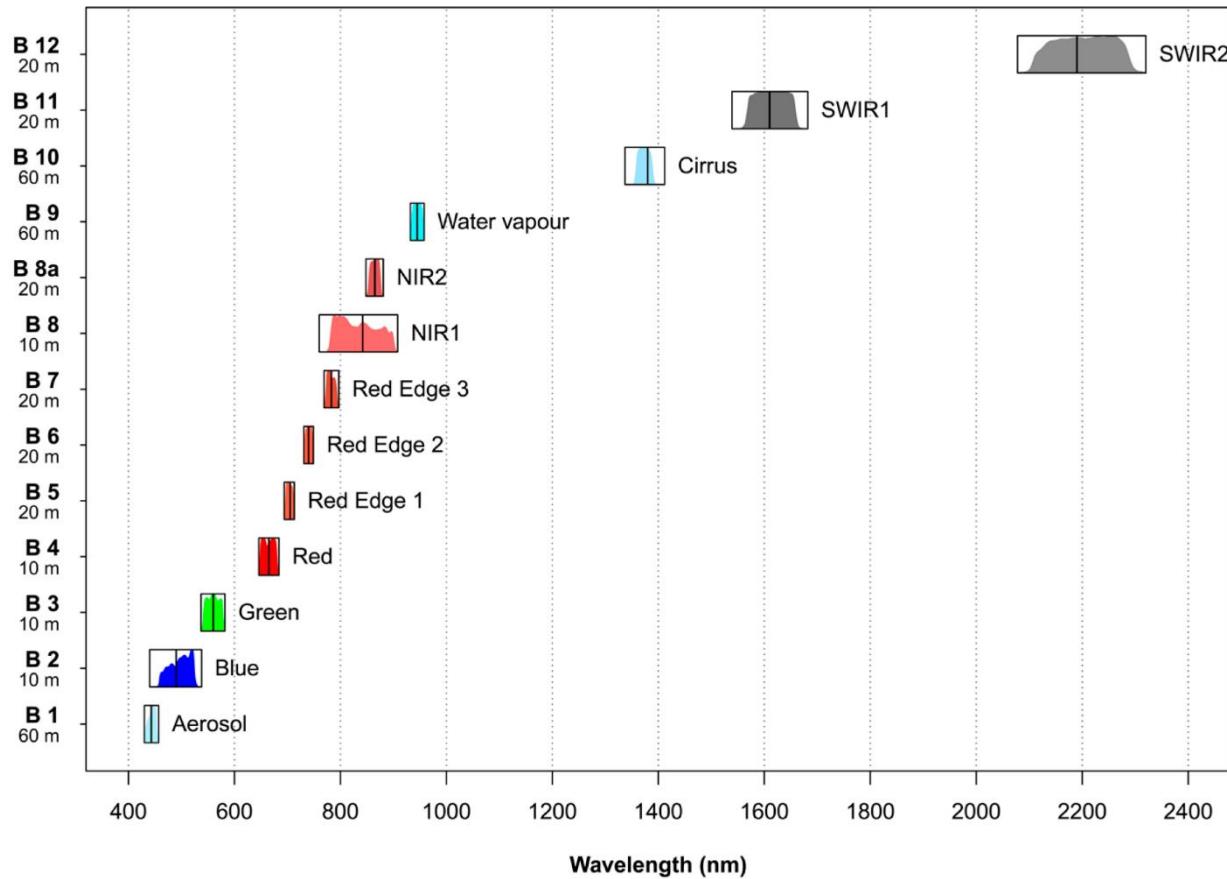
Remote Sensing: Landsat



Landsat sensors, spectral channels and band-passes, superimposed on atmospheric transmission percentage (grey background). MSS: Landsat-1 through -5; TM: Landsat-4 and -5; ETM+: Landsat-7; OLI and TIRS: Landsat-8. (Source: NASA/Landsat Legacy Project Team and American Society for Photogrammetry and Remote Sensing). Wulder et al., 2018

Remote Sensing: Sentinel 2

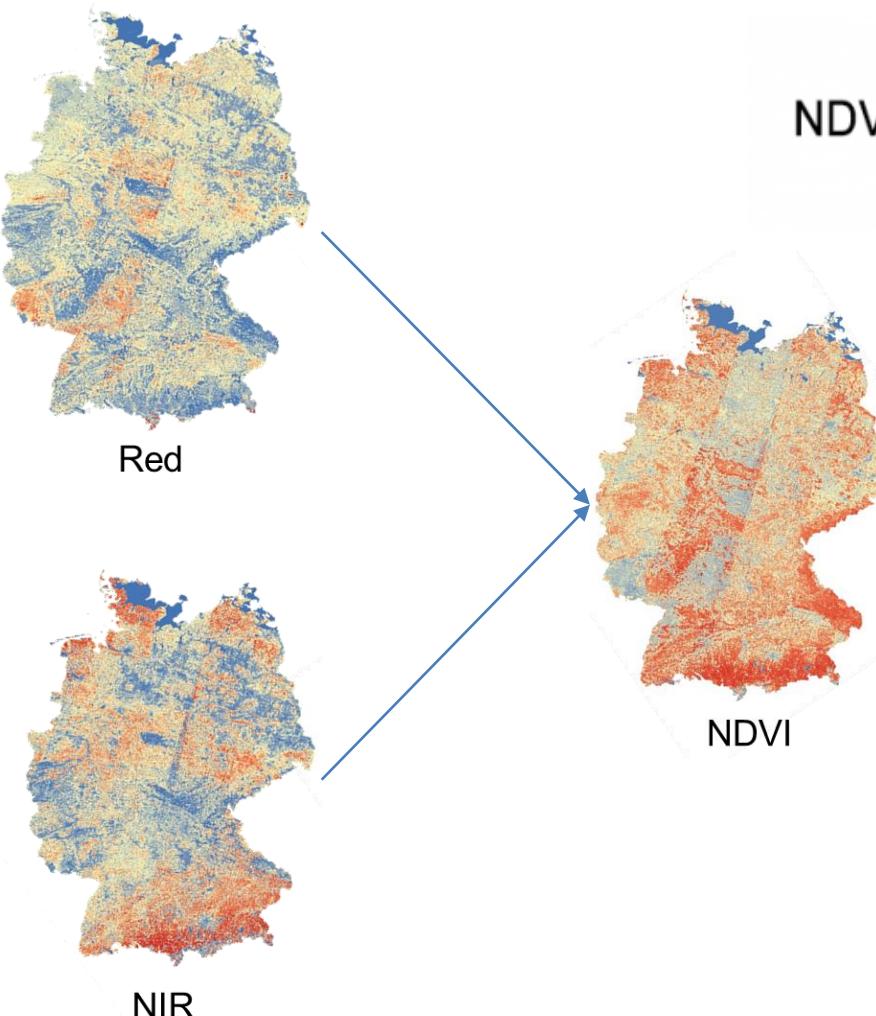
s
c
o
r
p
a
n



Characteristics of the Multi Spectral Instrument (MSI) on board Sentinel-2. The spectral response functions are shown in colors with the central wavelength in black. The band names and the corresponding spatial resolutions (in meters) are also indicated. Immitzer et al., 2016

Remote Sensing: band ratio

s
c
o
r
p
a
n

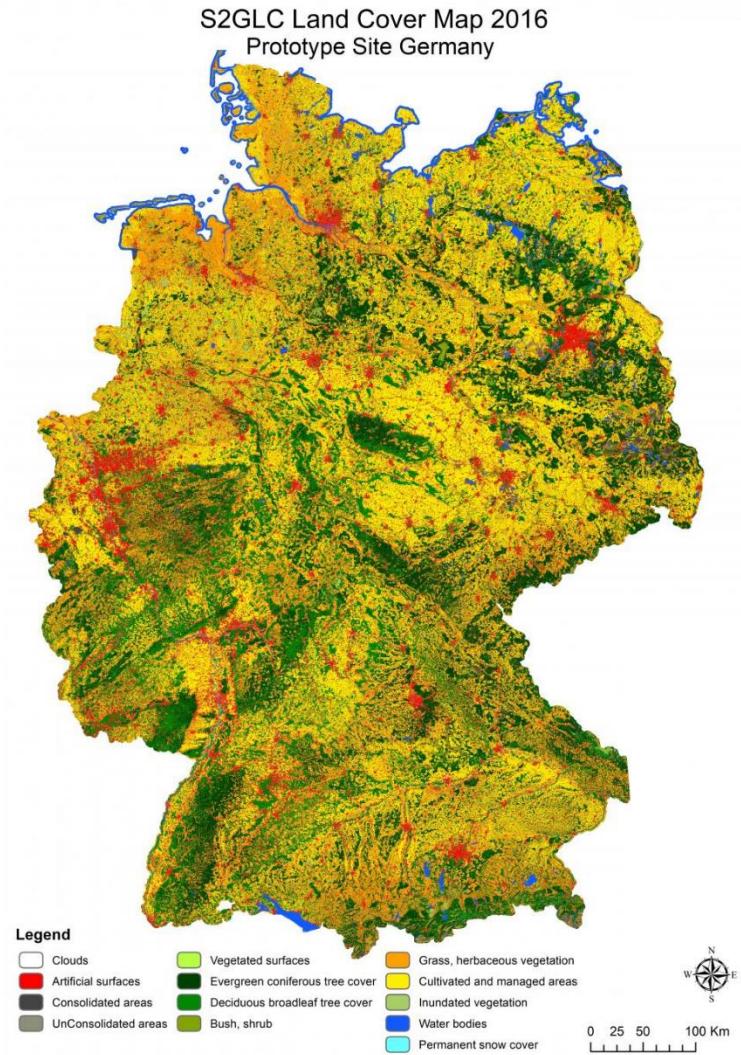


$$\text{NDVI} = \frac{(\text{NIR} - \text{Red})}{(\text{NIR} + \text{Red})}$$

Red and NIR bands of Landsat-8 for Germany

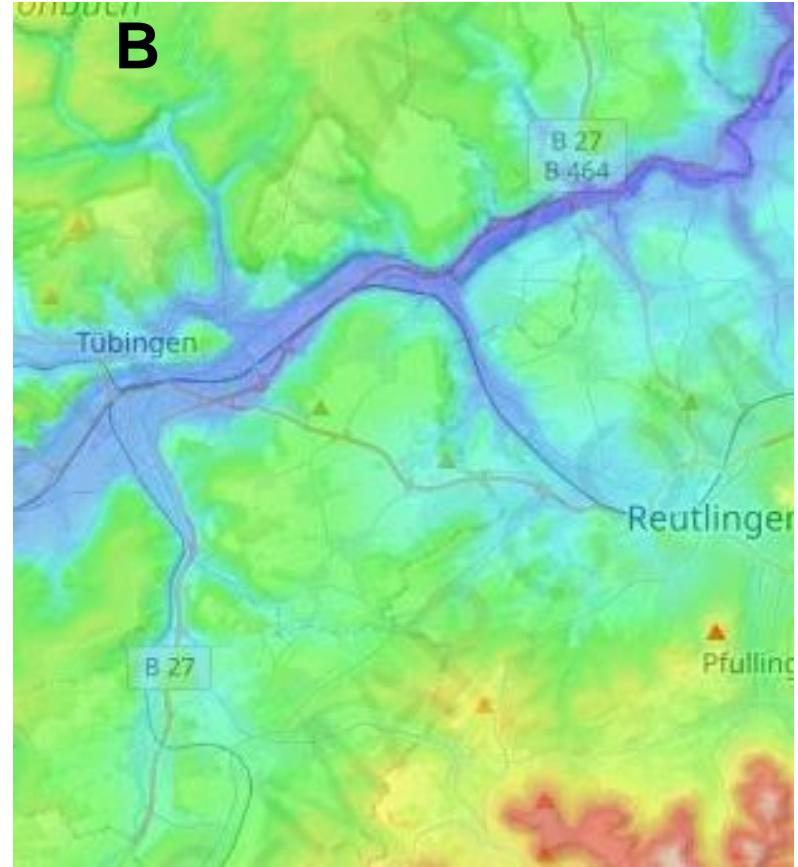
Remote Sensing: Land use/cover

S
C
O
R
P
A
N



Digital Elevation Model (DEM)

S
C
O
R
P
A
N

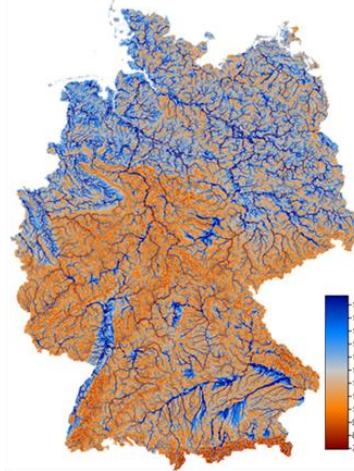


Elevation maps of Tuebingen, <https://en-gb.topographic-map.com/>

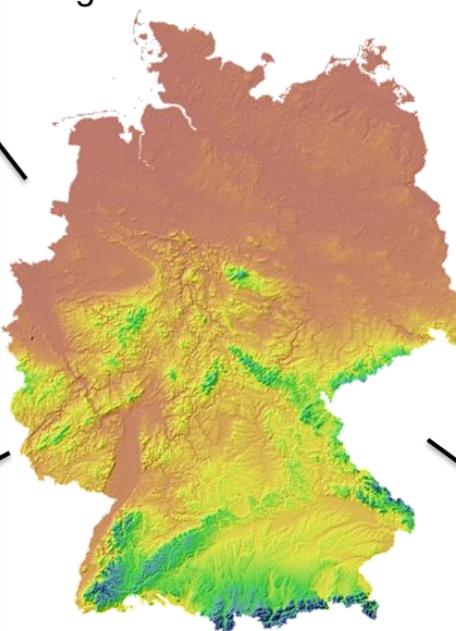
DEM: Terrain Attributes

S
C
O
r
p
a
n

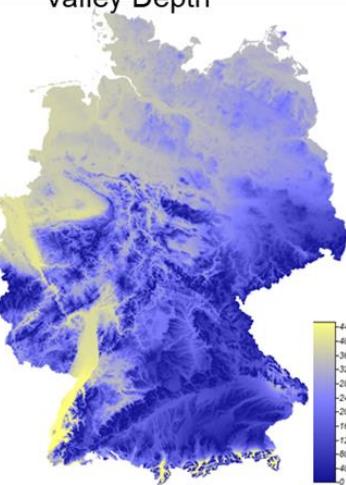
Topographic Wetness Index



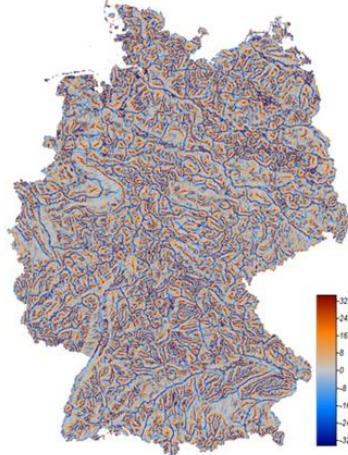
Digital Elevation Model



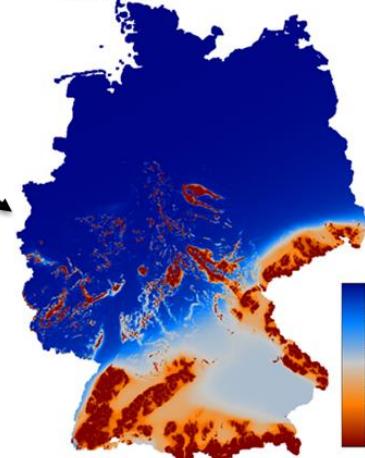
Valley Depth



Convergence Index



MrVBF

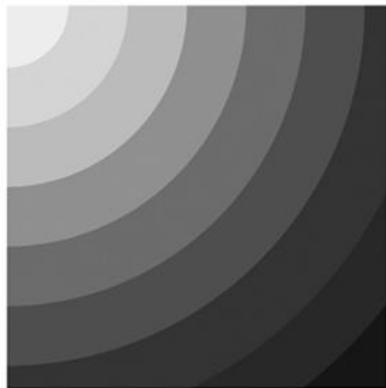


Basic Terrain Derivatives of Germany

Distance layers

S
C
O
R
P
A
N

Upper left (C1)



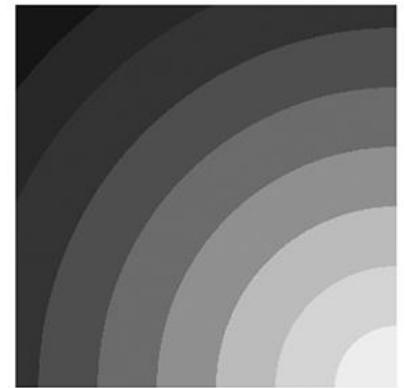
Upper right (C2)



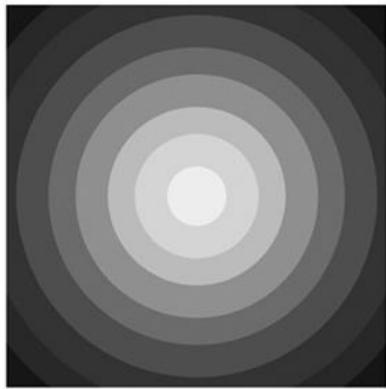
Lower left (C3)



Lower right (C4)



Centre (CC)



European Journal of Soil Science, 2018

doi: 10.1111/ejss.12687

Spatial modelling with Euclidean distance fields and machine learning

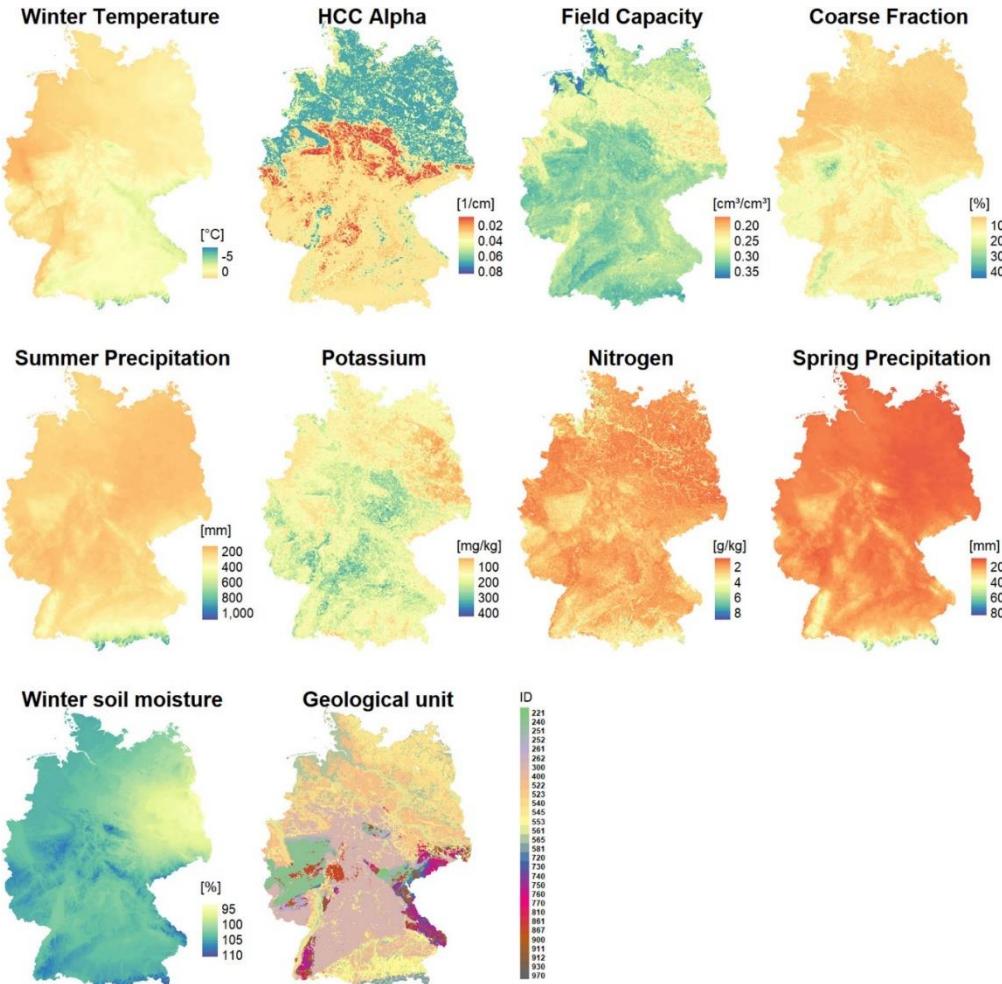
T. BEHRENS^a , K. SCHMIDT^a, R. A. VISCARRA ROSSEL^b , P. GRIES^a , T. SCHOLTEN^a & R. A. MACMILLAN^c

^aDepartment of Geosciences, Soil Science and Geomorphology, University of Tübingen, Rümelinstraße 19-23, 72074, Tübingen, Germany,

^bCSIRO Land and Water, GPO Box 1700, Canberra ACT 2601, Australia, and ^cLandMapper Environmental Solutions Inc., 7415 118 A Street NW, Edmonton, Alberta Canada

Other Covariates

s
c
o
r
p
a
n



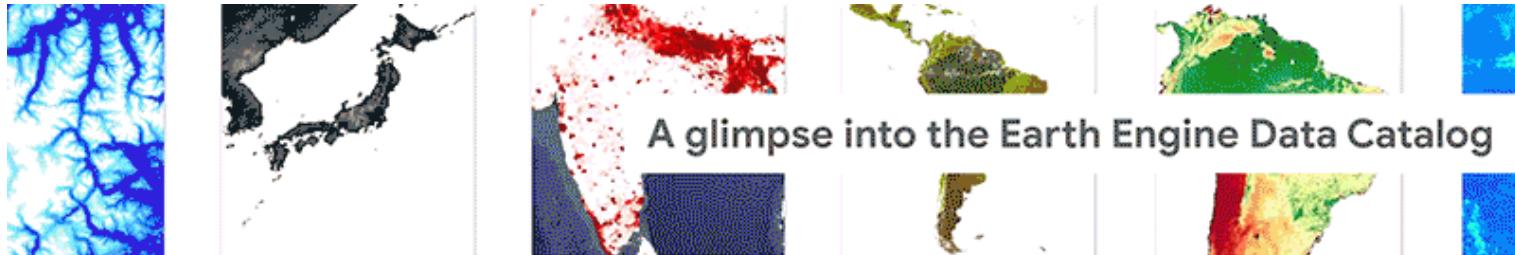
Other Geospatial Data of Germany (Petermann et al., 2020)

Where Can I Obtain the Covariates?

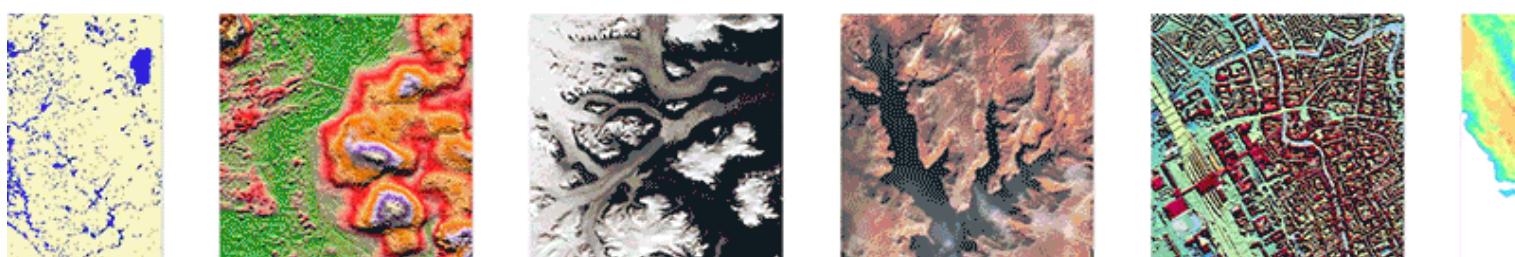
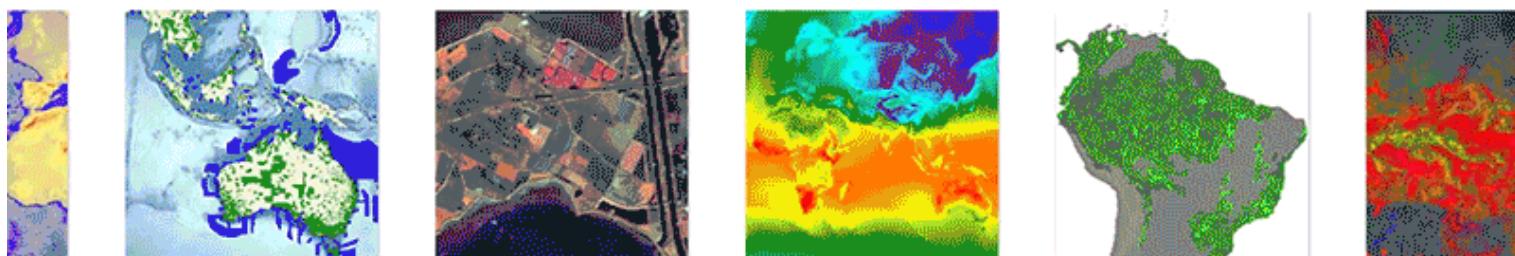
s
c
o
r
p
a
n



Google Earth Engine



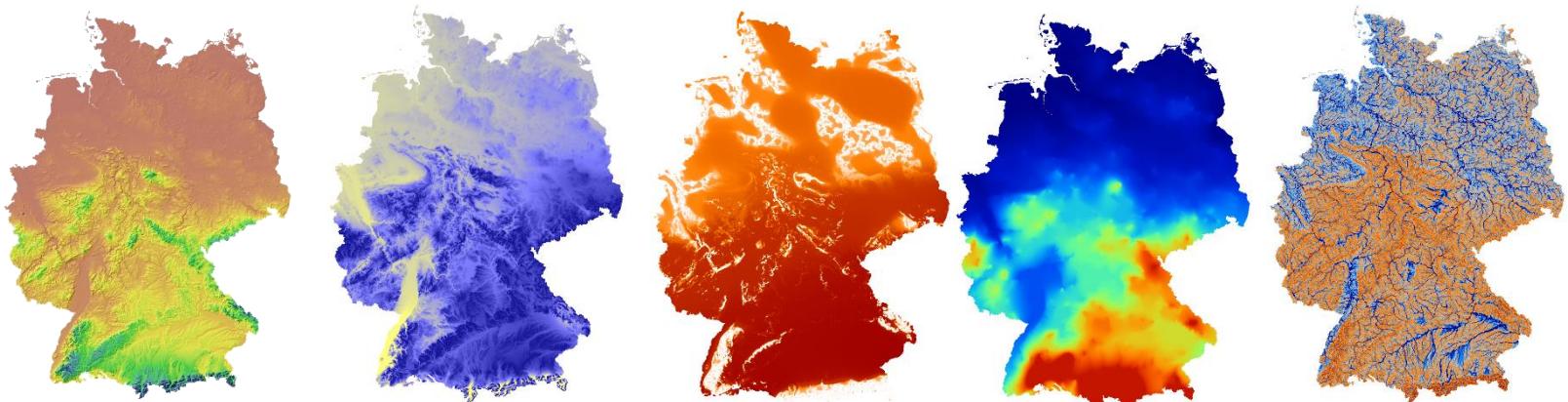
A glimpse into the Earth Engine Data Catalog



Choosing Covariates

s
c
o
r
p
a
n

- What factors influence soil formation?
 - Link to pedological knowledge
- What data do I have?
- Data should represent the multiple SORPAN covariates that influence soil development in the area

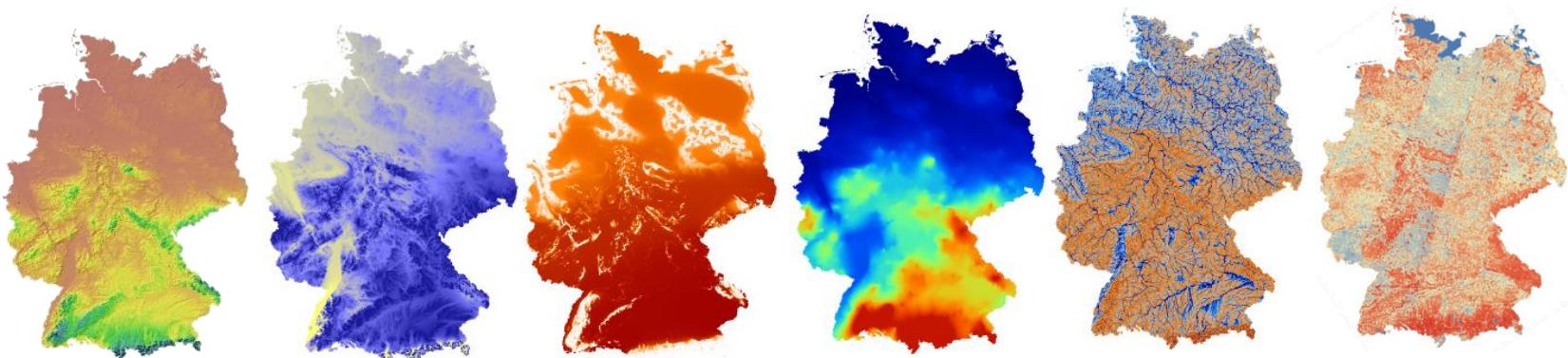


Technical and Practical Notes

s
c
o
r
p
a
n

Covariate data sources

- SRTM digital elevation model (**30 m resolution**)
- WorldDEM digital elevation model (**12 m resolution**)
- Landsat-8 satellite images (**30 m resolution**)
- Sentinel-2 satellite images (**10, 20, and 60 m resolution**)
- MODIS satellite images (**250, 500, and 1000 m resolution**)
- Global Land Cover maps (**30 m resolution**)
- JAXA's ALOS radar images (**20 m resolution**)
- Monthly precipitation images (**1000 m resolution**)
- Geology maps (**polygon**)



Technical and Practical Notes

s
c
o
r
p
a
n

Preparing covariate layers

1. Converting polygon maps to rasters
2. Downscaling or upscaling rasters to a common resolution,
3. Filtering out missing pixels/reducing noise and multicollinearity problems,
4. Overlaying raster stacks and points

Technical and Practical Notes

1. *Converting polygon maps to rasters*

s

c

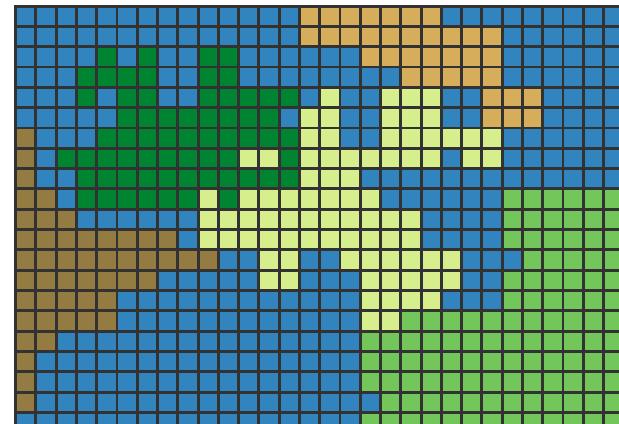
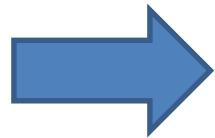
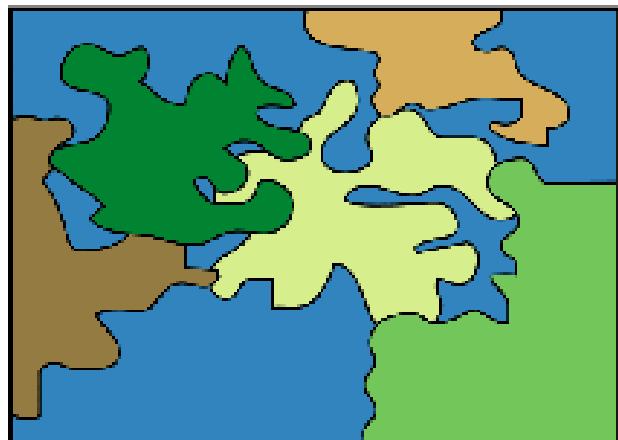
o

r

p

a

n

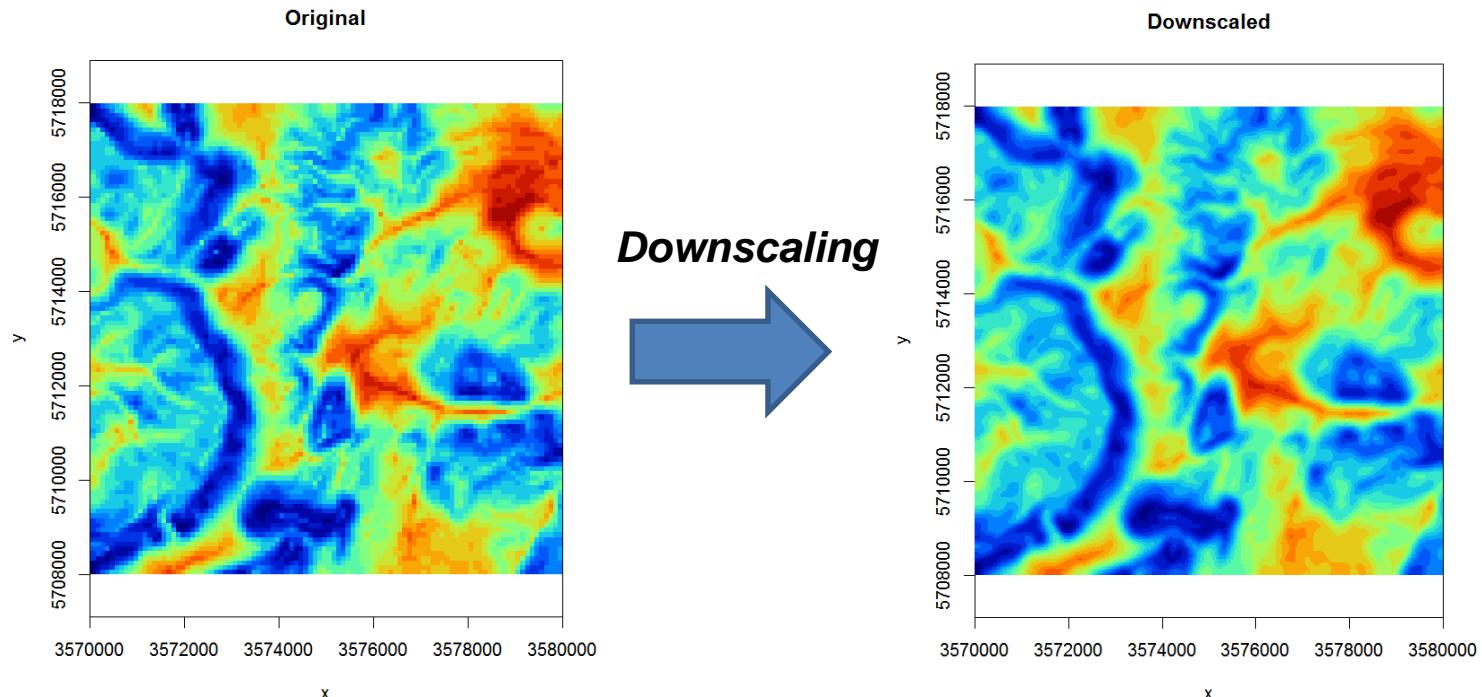


Technical and Practical Notes

s
c
o
r
p
a
n

2. Downscaling or upscaling rasters

- To adjust the resolution of some covariates that have either too coarse or too fine a resolution compared to the target resolution
- The process of bringing raster layers to a common grid resolution is also known as resampling

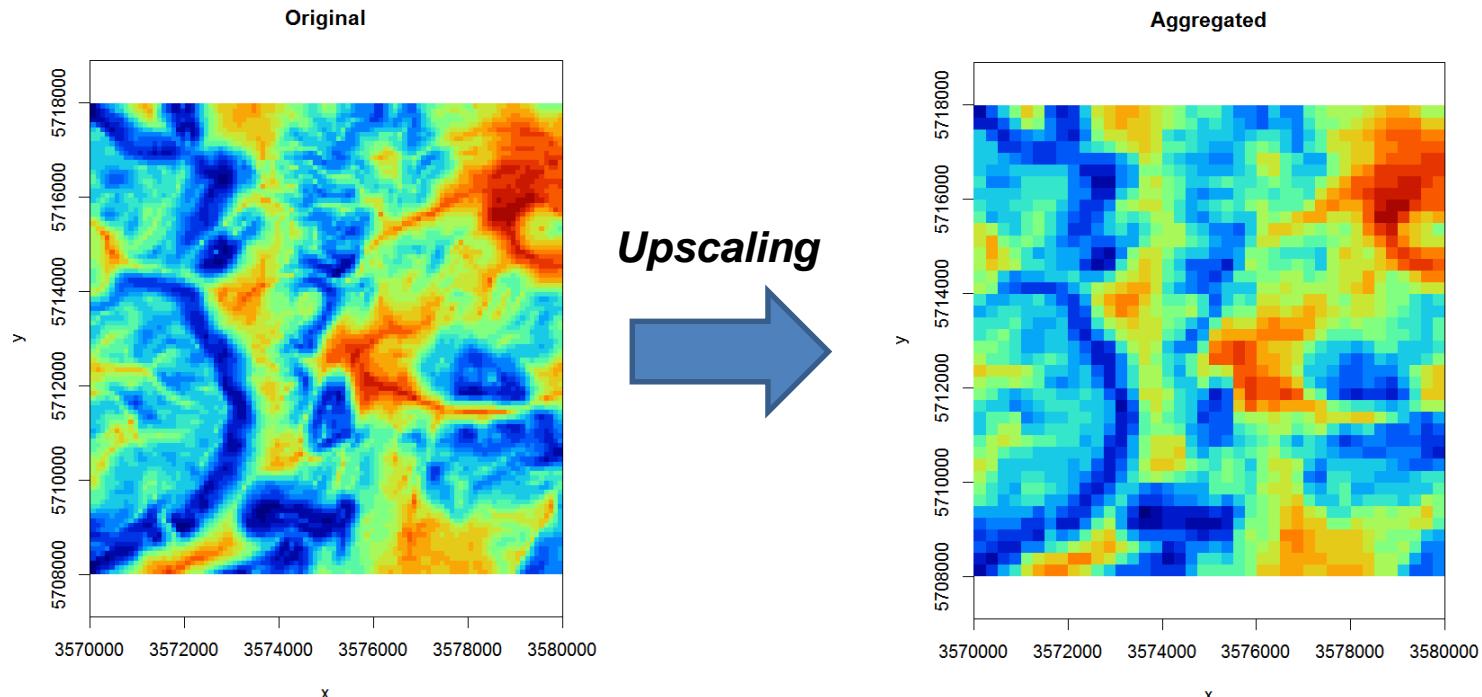


Technical and Practical Notes

s
c
o
r
p
a
n

2. Downscaling or upscaling rasters

- To adjust the resolution of some covariates that have either too coarse or too fine a resolution compared to the target resolution
- The process of bringing raster layers to a common grid resolution is also known as resampling

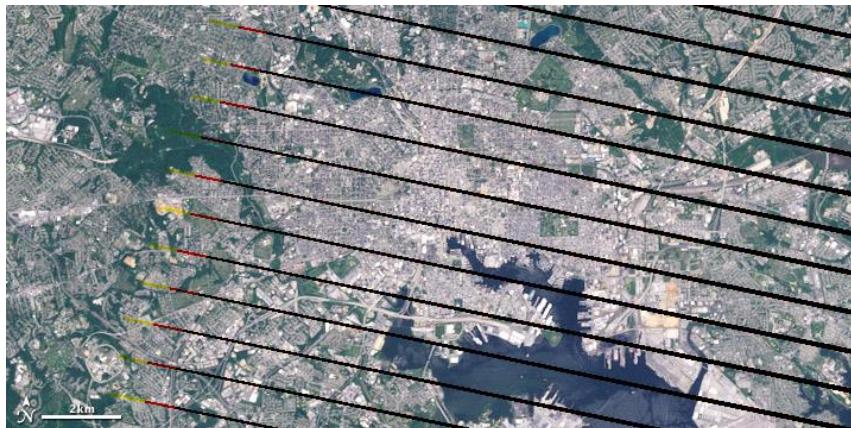


Technical and Practical Notes

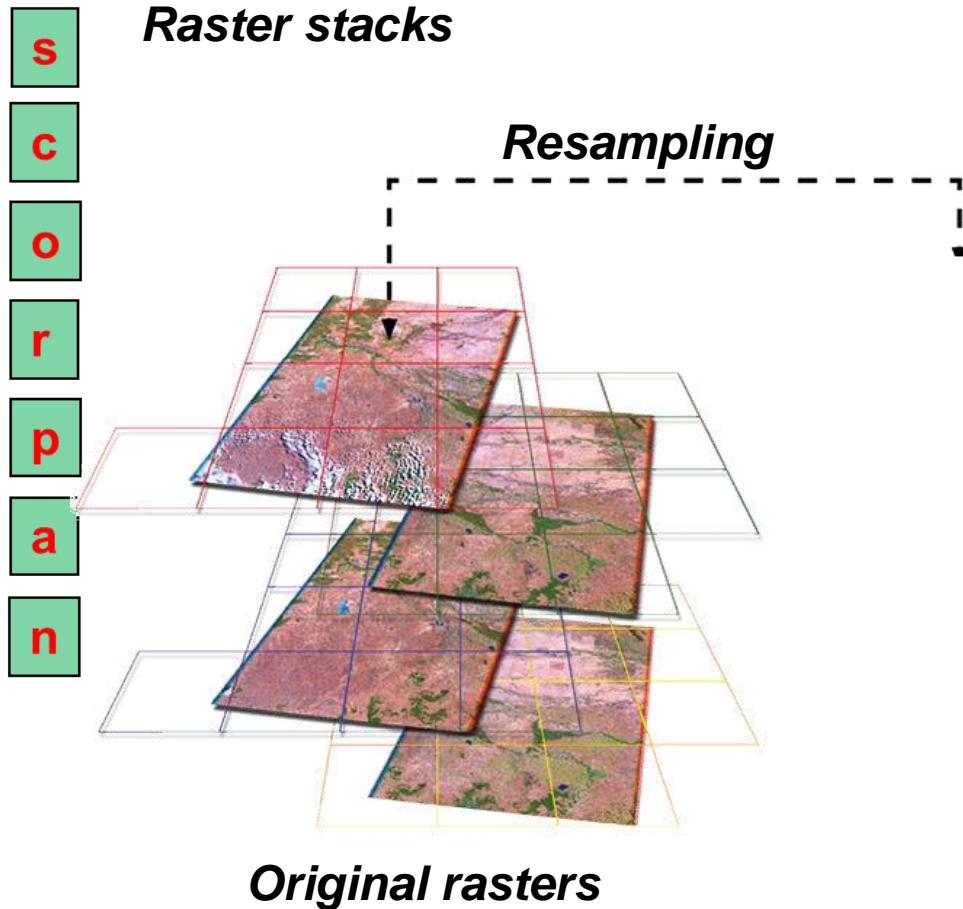
s
c
o
r
p
a
n

3. ***Filtering out missing pixels and artifacts***

- could cause serious problems for producing soil maps as the missing pixels and artifacts would propagate to predictions: if only one layer in the raster stack misses values then predictive models might drop whole rows in the predictions even though data is available for 95% of rows.
- Missing pixels can be efficiently filtered by using for example the gap filling functionality available in the SAGA GIS
- Another way to filter the missing pixels, to reduce noise and to reduce data overlap is to use Principal Components transformation of original data



Technical and Practical Notes

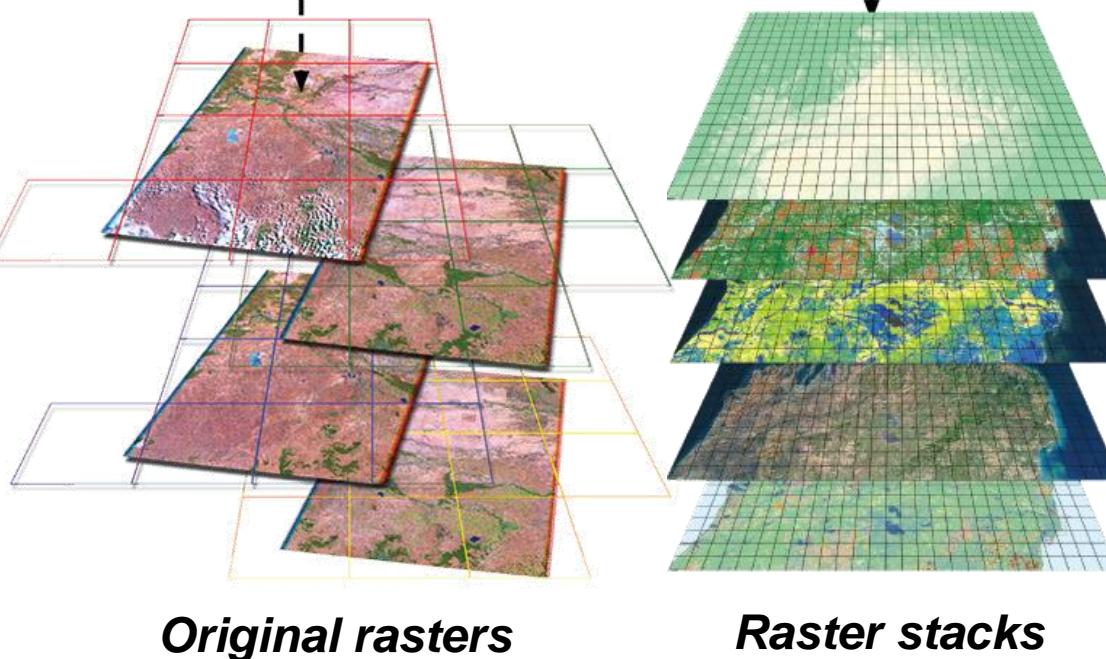


Technical and Practical Notes

s
c
o
r
p
a
n

Raster stacks

Resampling



Technical and Practical Notes

s

c

o

r

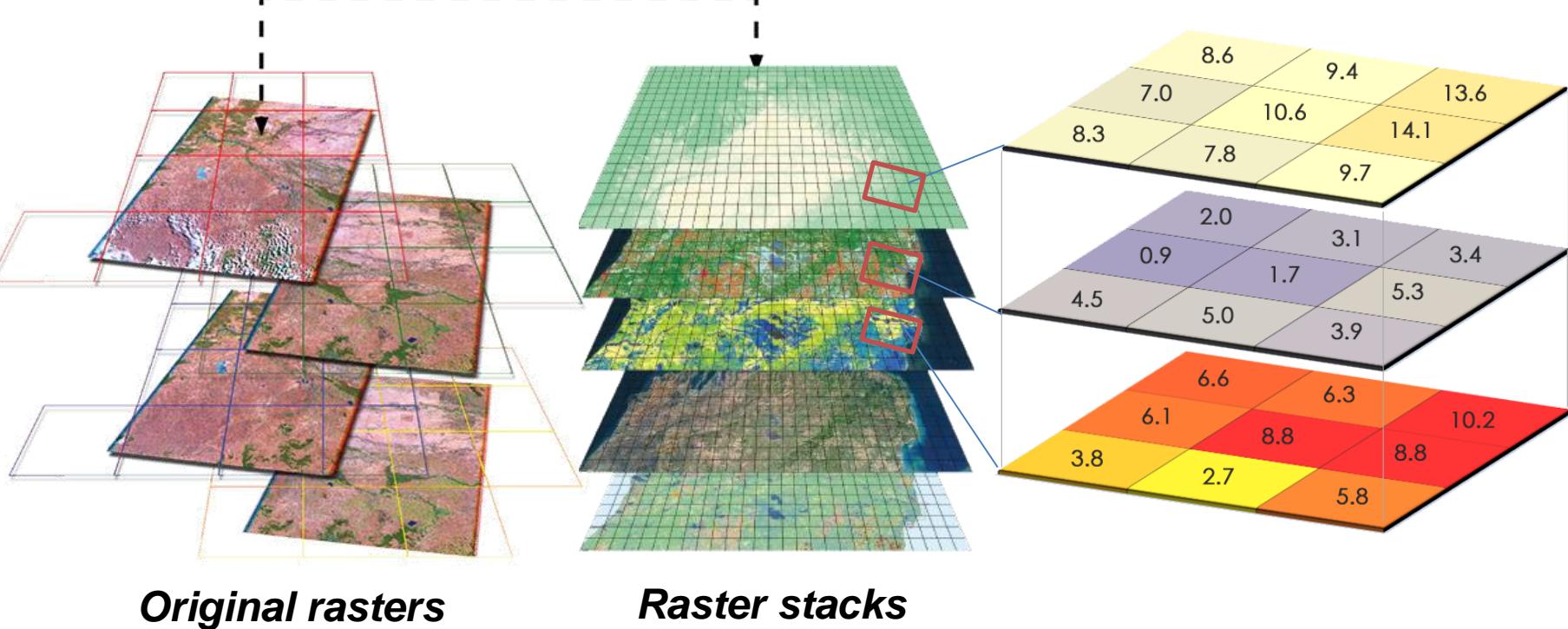
p

a

n

Raster stacks

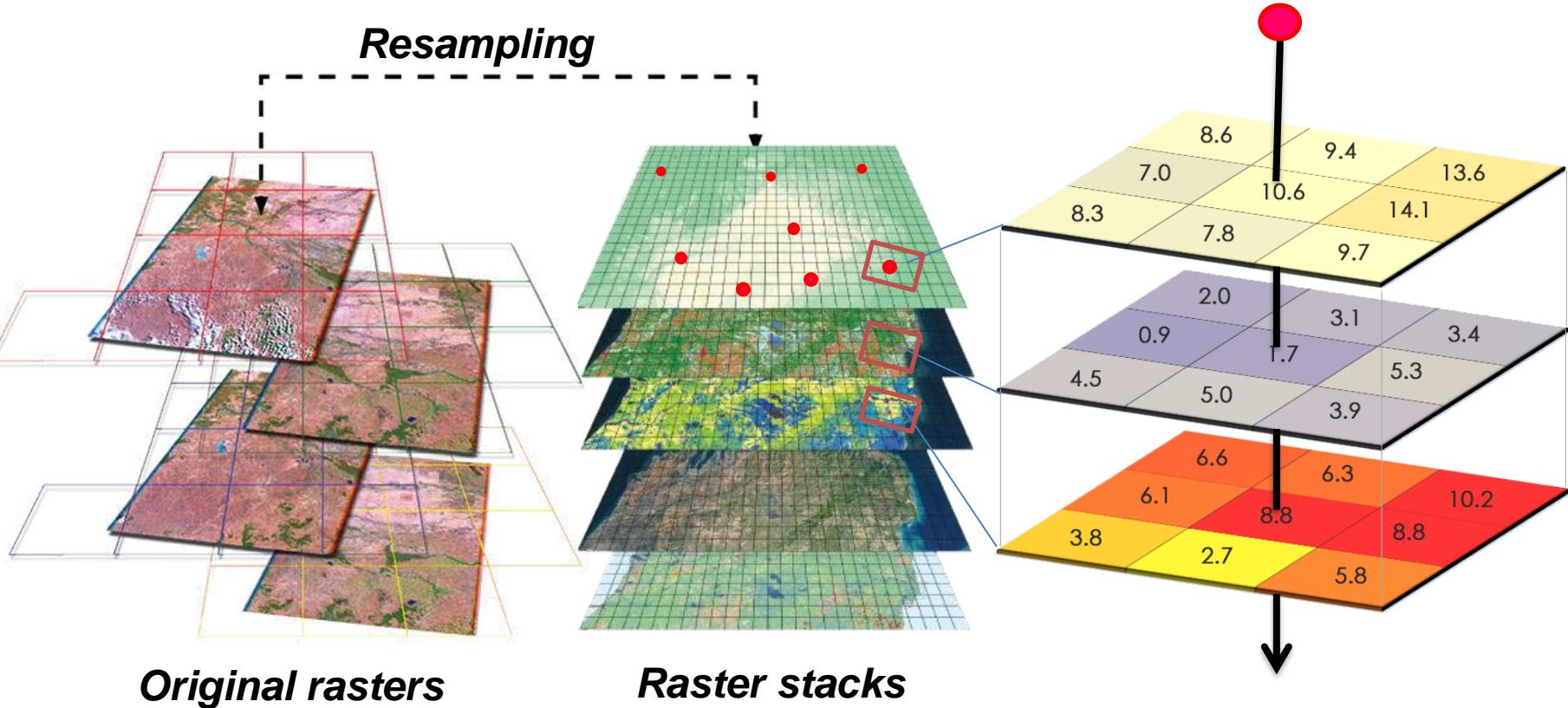
Resampling



Technical and Practical Notes

s
c
o
r
p
a
n

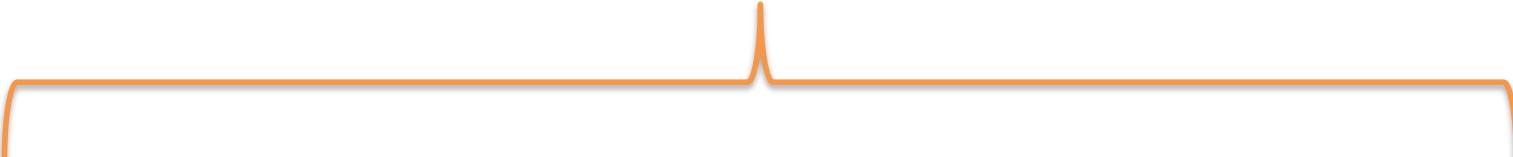
4. Overlaying raster stacks and points



Technical and Practical Notes

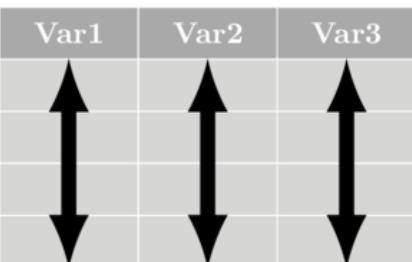
Geo-database
Tidy data
(*data.frame*)

Predictors
Covariates
Independent variables
Features



X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	Class
-0.60	3.06	-5.77	0.69	1.46	1.22	1.94	0.61	-0.07	-2.02	-0.87	0.71	0.89	-0.91	0.21	a
0.61	1.61	0.21	2.74	0.46	1.29	0.13	0.55	-0.60	-0.53	0.70	-0.41	-0.62	1.17	-0.32	b
3.81	0.99	-1.21	0.30	0.11	-0.34	-0.29	0.15	-0.09	1.30	0.08	0.66	-0.44	0.45	-0.13	a
0.50	0.14	-0.30	0.64	-1.28	0.28	0.76	-0.61	0.10	-0.07	0.41	-0.53	-0.62	0.06	-0.52	c
1.46	0.23	0.17	0.53	-0.94	0.37	0.64	-0.83	0.20	-0.06	0.18	-0.35	-0.64	0.39	-0.22	a
3.07	-0.19	1.31	0.84	-0.66	1.05	0.94	-1.12	-0.23	-0.31	0.45	-0.54	-0.16	0.38	0.53	a

Variables in Columns

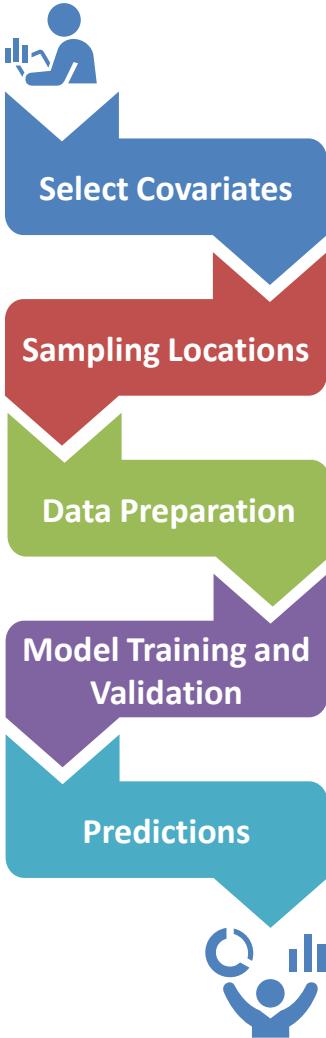


Observations in Rows



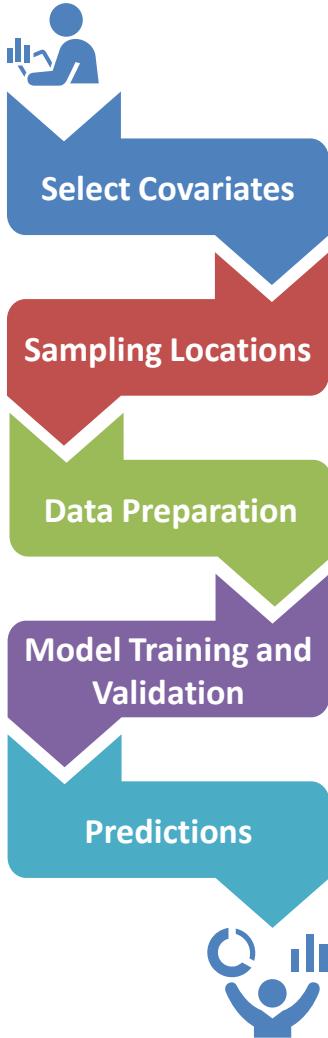
Response variable
Target variable
Dependent variable
(Soil clay or soil types)

Take-Home Message



- 1 **Environmental covariates**, relevant as predictors of soil property/class, are derived from remote sensing, digital elevation, climatic datasets, ...
- 2 **Soil samples** are collected at the specified locations (e.g., Latin hypercube sampling) and soil property is measured in the laboratory.
- 3 Intersecting the covariates with the soil point observations.
- 4 Machine learning models (e.g., random forest) are trained using training data, and accuracy assessment is carried out using the test data set.
- 5 The ML models are applied to the entire study area in order to produce a **soil property/class map**.

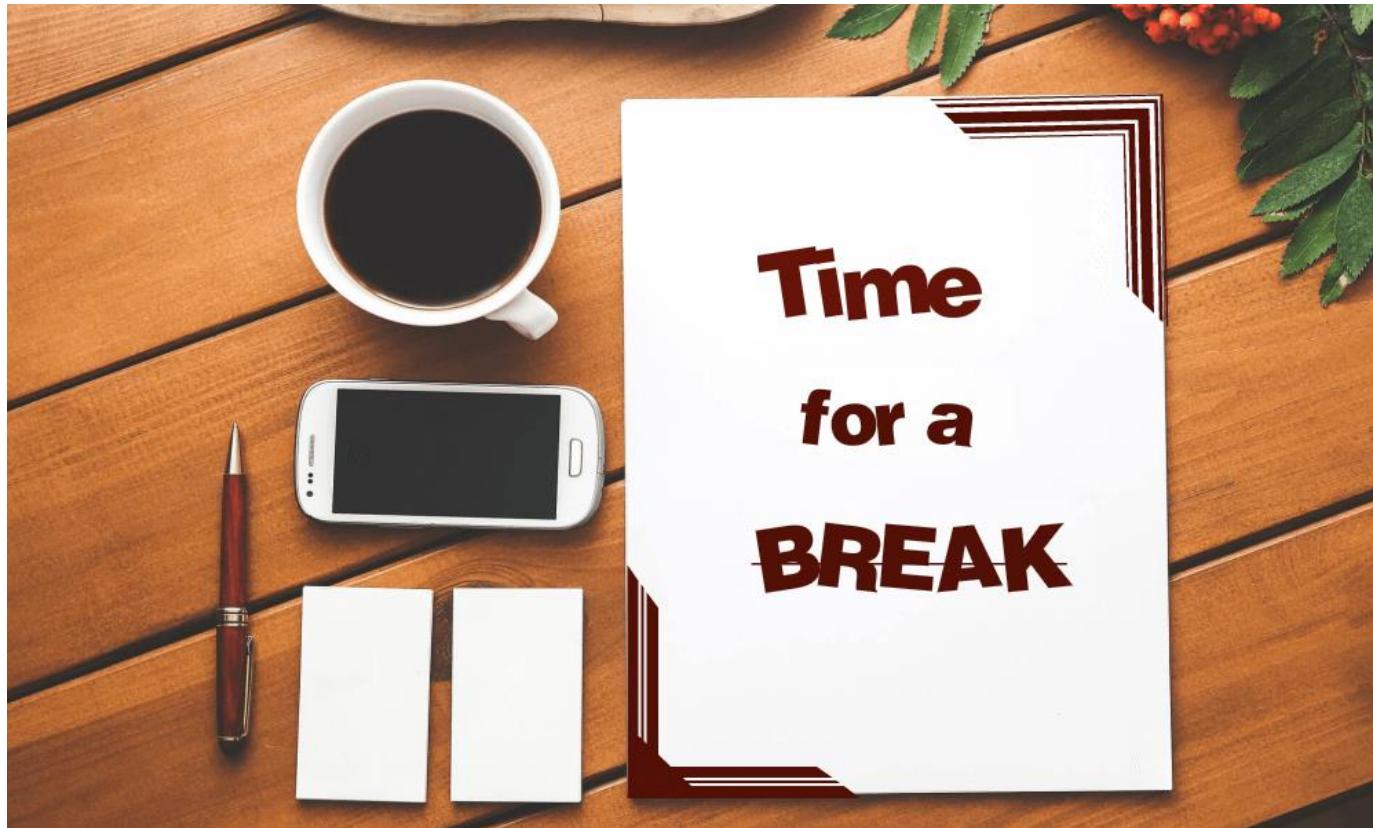
Take-Home Message



- 1 **Environmental covariates**, relevant as predictors of soil property/class, are derived from remote sensing, digital elevation, climatic datasets, ...
- 2 **Soil samples** are collected at the specified locations (e.g., Latin hypercube sampling) and soil property is measured in the laboratory.
- 3 Intersecting the covariates with the soil point observations.
- 4 Machine learning models (e.g., random forest) are trained using training data, and accuracy assessment is carried out using the test data set.
- 5 The ML models are applied to the entire study area in order to produce a **soil property/class map**.

Please install R and RStudio

Please install SAGA and Google Earth



1. Open Rstudio

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with various icons. A script editor window titled "Untitled1*" contains the following R code:

```
1 1 + 1
2 2 + 3 + 4
3 x <- c(1:100)
4 hist(x)
5
```

The console window displays the standard R startup message, which includes:

You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

A search bar at the bottom says "Type here to search".

2. Clear Project List

The screenshot shows the RStudio interface with the Environment tab selected. A red arrow points from the "Environment" tab in the left panel to the "Project List" option in the context menu. The context menu is open and lists the following options:

- New Project...
- Open Project...
- Open Project in New Session...
- Close Project
- Clear Project List** (highlighted)
- Project Options...

A gray box labeled "Environment is" is positioned near the context menu.

3. New Project

The screenshot shows the RStudio interface. In the top right corner, a red arrow points to the 'Project' dropdown menu, which is open to show options like 'New Project...', 'Open Project...', and 'Clear Project List'. The 'Environment' tab is selected in the main workspace. On the left, there's an 'Untitled1*' script editor with some R code. Below it is the 'Console' window displaying the standard R welcome message. The bottom status bar shows the date and time.

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Untitled1* Go to file/function Addins

1 1 + 1
2 2 + 3 + 4
3 x <- c(1:100)
4 hist(x)
5

5:1 (Top Level) R Script

Console ~/
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
Natural language support but running in an English locale
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

Type here to search

Project: (None)

New Project...
Open Project...
Open Project in New Session...
Close Project
Clear Project List
Project Options...

Environment History Connections

Import Dataset

Global Environment

Files Plots Packages Help Viewer

4:35 PM 2/9/2021

4. Existing Directory

The screenshot shows the RStudio interface with a 'Create Project' dialog box overlaid. The dialog has three options: 'New Directory', 'Existing Directory', and 'Version Control'. A large red arrow points from the top right towards the 'Existing Directory' option. The RStudio environment includes a code editor with some R code, a console window displaying a welcome message, and a global environment viewer.

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Untitled1* Go to file/function Addins

Project: (None)

1 1 + 1
2 2 + 3 + 4
3 x <- c(1:100)
4 hist(x)
5

5:1 (Top Level) ◆

Console ~/ ◆
You are welcome to redistribute
Type 'license()' or 'licence()'
Natural language support but r
R is a collaborative project wit
Type 'contributors()' for more i
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |

New Project

Create Project

New Directory Start a project in a brand new working directory >

Existing Directory Associate a project with an existing working directory > ↓

Version Control Checkout a project from a version control repository >

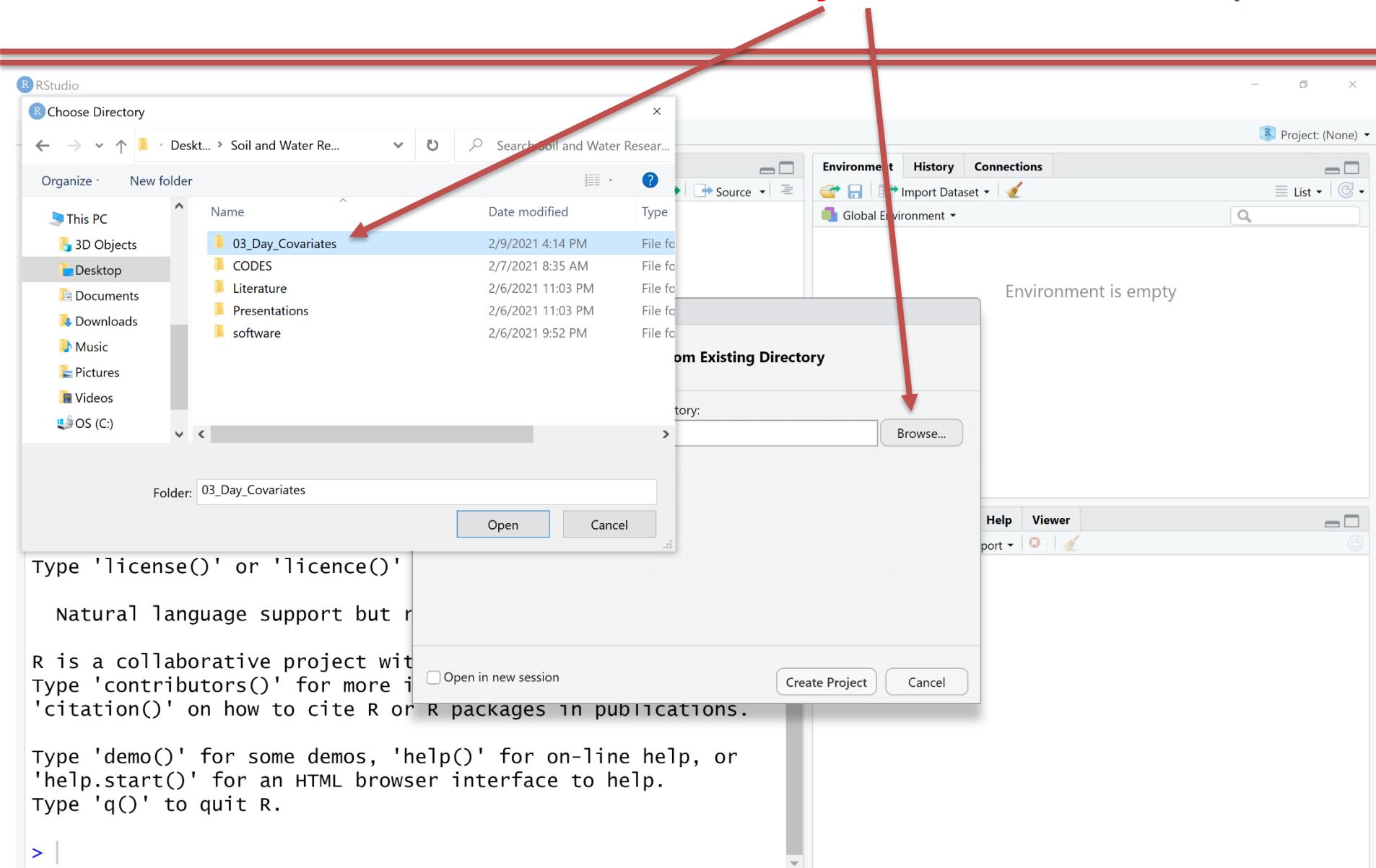
Environment History Connections

Global Environment

Environment is empty

Help Viewer

5. Find the folder and find **03_Day_Covariates** and open



6. Create project

The screenshot shows the RStudio interface with a red arrow pointing to the 'Create Project' button in the 'Create Project from Existing Directory' dialog box.

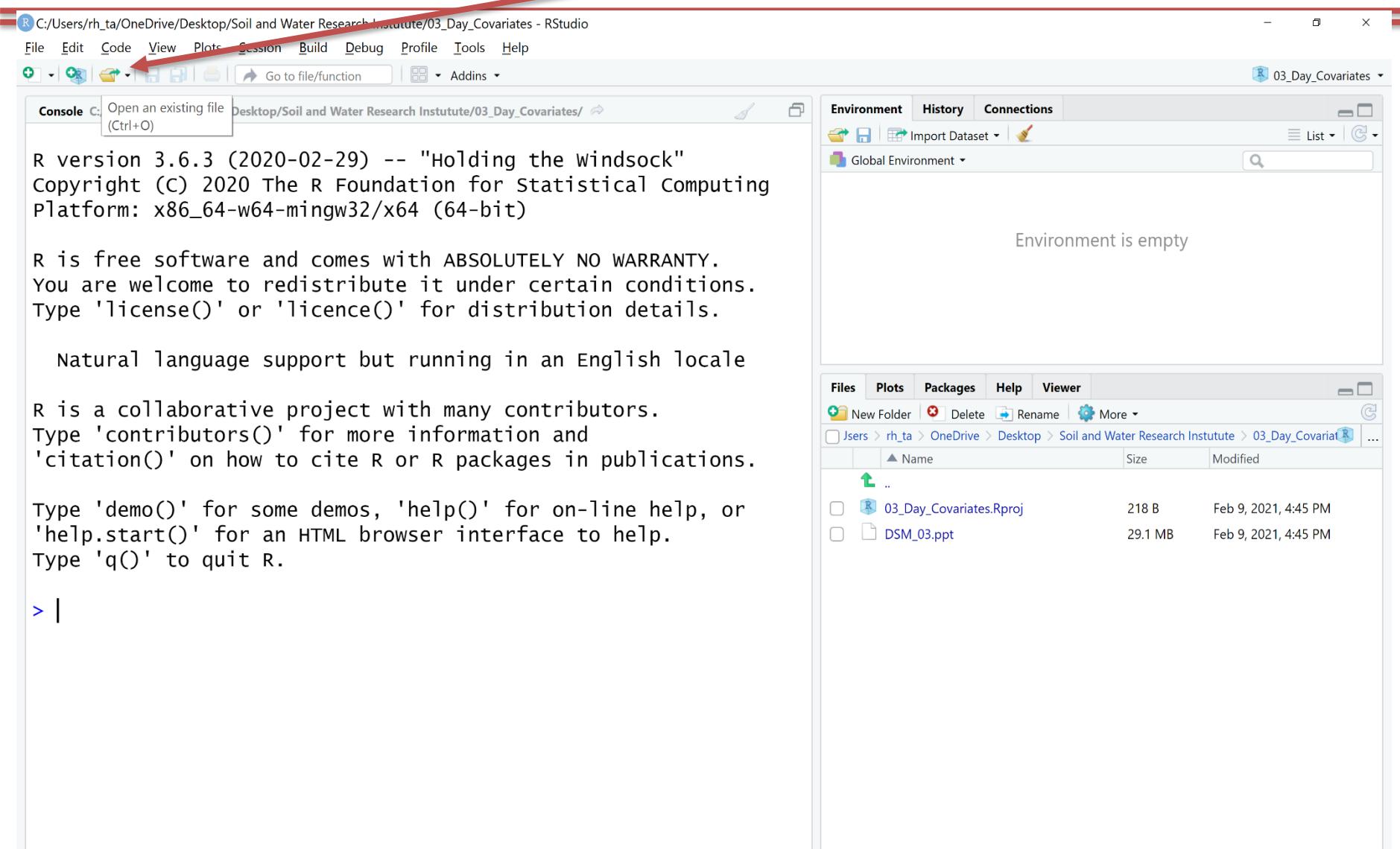
RStudio interface elements:

- File Edit Code View Plots Session Build Debug Profile Tools Help
- Addins
- Untitled1*
- Source on Save
- Run
- Environment History Connections
- Global Environment
- Environment is empty
- Console
- You are welcome to redistribute Type 'license()' or 'licence()'
- Natural language support but r
- R is a collaborative project with Type 'contributors()' for more information 'citation()' on how to cite R or R packages in publications.
- Type 'demo()' for some demos, 'help()' for on-line help, or 'help.start()' for an HTML browser interface to help.
- Type 'q()' to quit R.
- >

New Project dialog box details:

- Create Project from Existing Directory
- Project working directory: C:/Users/rh_ta/OneDrive/Desktop/Oil and Water R
- Back
- Browse...
- Open in new session
- Create Project
- Cancel

7. Open an existing file and find **03_R_Cov.R**



R C:/Users/rh_ta/OneDrive/Desktop/Soil and Water Research Institute/03_Day_Covariates - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Open an existing file Desktop/Soil and Water Research Institute/03_Day_Covariates/

R version 3.6.3 (2020-02-29) -- "Holding the Windsock"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |

Environment History Connections

Import Dataset

Global Environment

Environment is empty

Files Plots Packages Help Viewer

New Folder Delete Rename More

Name	Size	Modified
03_Day_Covariates.Rproj	218 B	Feb 9, 2021, 4:45 PM
DSM_03.ppt	29.1 MB	Feb 9, 2021, 4:45 PM

8. Open 03_R_Cov.R

